Arrival Management Support in the Presence of Prediction Uncertainty

Tielrooij, M.

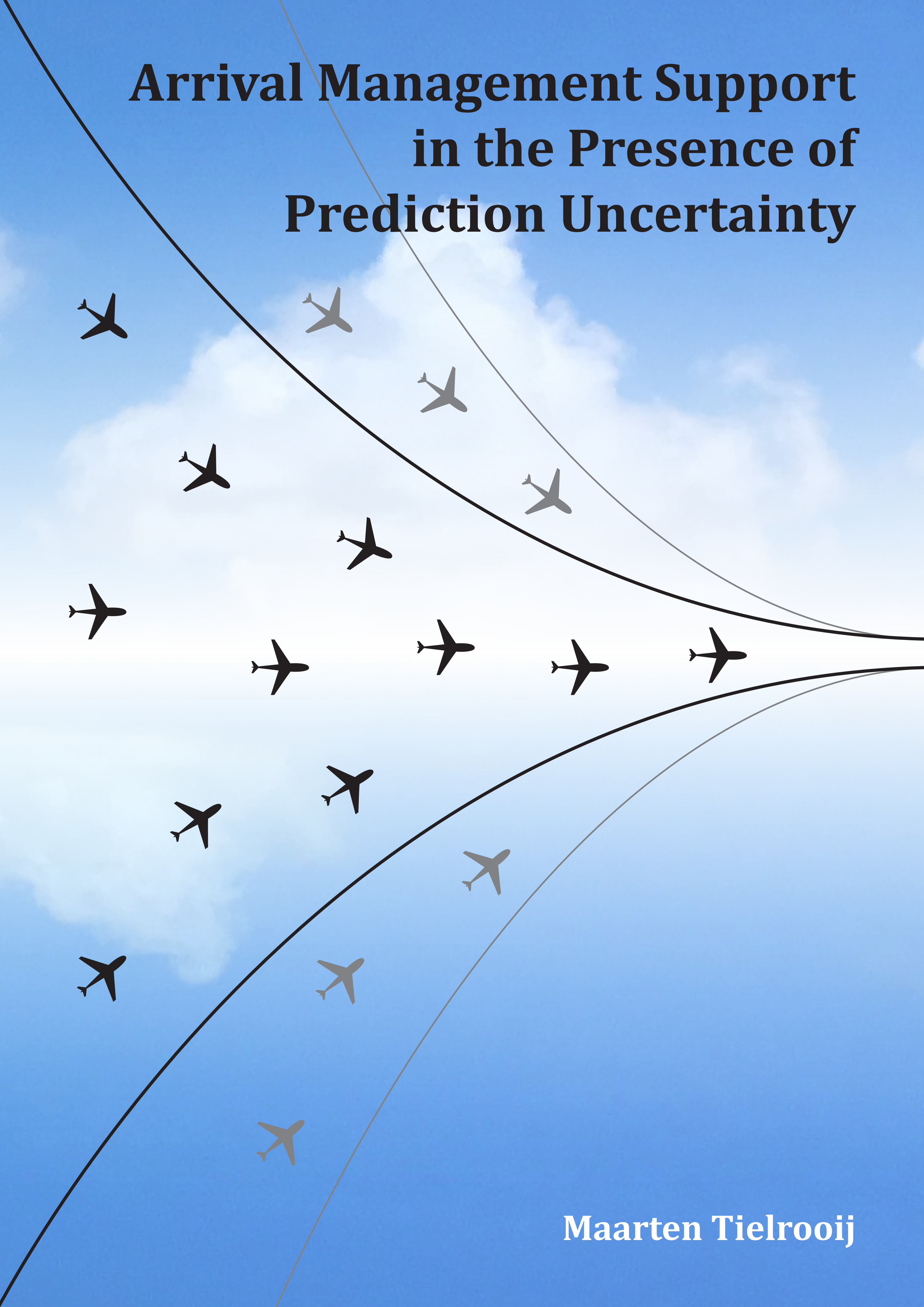
DOI
10.4233/uuid:a403112b-48a2-40d7-9db3-a5b754f31eee

Publication date
2022

Document Version
Final published version

Citation (APA)
Tielrooij, M. (2022). *Arrival Management Support in the Presence of Prediction Uncertainty*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:a403112b-48a2-40d7-9db3-a5b754f31eee

Arrival Management Support in the Presence of Prediction Uncertainty



Maarten Tielrooij

Arrival Management Support in the Presence of Prediction Uncertainty

Arrival Management Support in the Presence of Prediction Uncertainty

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus prof. dr. ir. T.H.J.J. van der Hagen
chair of the Board for Doctorates
to be defended publicly on
Thursday 20 October 2022 at 12:30 hours

by

Maarten TIELROOIJ

Master of Science in Aerospace Engineering,
Delft University of Technology, the Netherlands
born in Groningen, the Netherlands

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof. dr. ir. M. Mulder,	Delft University of Technology, promotor
Dr. ir. M.M. van Paassen,	Delft University of Technology, promotor
Independent members:	
Prof. dr. J. Lundberg MA	Linköping U., Sweden
Prof. dr. -ing. H. Vallery	Delft University of Technology
Dr. B.G. Hilburn	CHPR, USA
Prof. dr. -ing. H. Fricke	TU. Dresden, Germany
Prof. dr. ir. J.M. Hoekstra	Delft University of Technology

This dissertation is financed by Knowledge Development Centre Mainport Schiphol (KDC) as part of the AMAN cluster.

Keywords: Arrival Management, Trajectory Prediction, Prediction Uncertainty, Air Traffic Control, Demand and Capacity Balancing

Printed by: Ipskamp Printing with support from BzAc

Front & Back: Certain and uncertain arrivals in a single flow by Chris and Diana

ISBN 978-94-6421-878-7

An electronic version of this dissertation is available at
`http://repository.tudelft.nl/`.

Summary

Future concepts for Air Traffic Management (ATM) foresee—and even mandate—an increase in the distance from the destination airport at which the ATM system should start planning the arrival sequence and arrival time of inbound flights. By increasing the horizon beyond one hour before arrival, it becomes possible to timely adjust the flight path of the inbound aircraft to meet the planned time. These timely adjustments tend to result in a more efficient flight profile, reducing fuel burn and emissions.

However, the current prediction capabilities—be it on the ground or aboard the aircraft—do not yet provide the accuracy that would be required in arrival management. Even onboard Flight Management Systems can only achieve sufficient accuracy at about 30 minutes before arrival. Beyond that horizon, the uncertainty of the predicted Estimated Time of Arrival (ETA) prevents a decision on the landing sequence or the arrival times with a sufficiently low risk of revision. Such a revision of an arrival plan requires new adjustments to the flight's trajectory, which may void any efficiency benefit of an early decision or even lead to a net loss of efficiency.

Improved algorithms, better information sharing, and more computing power may reduce prediction uncertainty. But, beyond the prediction horizon of one hour, more and more processes may disturb the predicted trajectory. This is especially true in Europe, where many larger airports are inside this prediction horizon. It is unlikely that the processes on the ground can be predicted up to the accuracy required for arrival management.

Therefore, if the arrival management horizon is to be increased, support has to be developed for human operators or automated systems to perform arrival management in the presence of prediction uncertainty. This dissertation aims to develop means to work with the uncertainty by predicting it and developing a technique to visualise it to support a human operator as a sequence manager or the manager of an automated arrival management system.

By analysing historic prediction errors, predicting the uncertainty of a predicted arrival time is possible. Such a forecast allows online determination of the uncertainty per flight. The empirical method applies to predictions irrespective of the

origin of the prediction. Therefore, it works in a scenario where another Air Navigation Service Provider or the aircraft provides the prediction. By clustering the predictions on the properties of the flight, the uncertainty of an ETA is described using the Johnson distribution. Since the distribution parameters are tabulated per class of prediction, the method is not computation-intensive and would work in an online environment. Furthermore, the method is independent of the method used to predict the trajectory and can be adjusted as and when better prediction information is available.

Extending the ubiquitous time line diagram for arrival management allows visualising the mathematical construct of uncertainty of the predicted arrival time. The diagram is extended by visually showing the relation between ETA, predicted demand on a runway, and the capacity of that runway. A further extension toward multiple runways allows showing the trade-off between *demand* (the number of inbound flights) and *capacity* (the number of runways made available for arriving aircraft).

Neither two human-in-the-loop experiments with a single time line nor an experiment with the second concept showed an advantage of visualising uncertainty. Participants indicated difficulty in understanding the concept of uncertainty as presented on the display. A fundamental problem here is the complexity of the work domain itself when uncertainty is introduced. In systems without uncertainty, the operator has to develop a mental model of the system to forecast how it *will* evolve given its current state. When uncertainty is introduced, the operator has to develop a different type of mental model: How it *could* evolve and how *likely* that is.

The experiments used a theoretical uncertainty. A synthesis of the display concept with the historic prediction uncertainty—from 2014—showed that the display would provide little useful information beyond a 30-minute horizon due to the large uncertainty in the predictions. With that dataset, the display would not be able to support arrival management over a longer horizon. Using more recent data or other data sources would improve the accuracy and therefore increase the useful horizon.

The synthesis of the display and the data does prompt a novel information item and a way to calculate that item: The stable sequence horizon. This parameter is the horizon within which it is unlikely that flights will change position in the sequence. This is an indicator of the risk that decisions on the sequence and on arrival times need revision. Such an indicator could support dynamic adjustment of the arrival management horizon, allowing timely decisions when warranted by the prediction uncertainty at that moment.

Contents

Chapter 1

Introduction

With a global profit margin of 6.9% in 2019 [1, p. 15], and the cost of fuel representing 24% [1, p. 32] of operating costs, airlines are continuously searching to optimise their flight profile. At the same time, high traffic volumes often require Air Traffic Control (ATC) to plan when flights can use the limited capacity at airports and in airspaces.

When flights are inbound to an airport, a conflict of interest exists between the goals of the single Airspace User (AU) and the goals of ATM: The objective of the aircraft operator is to fly a trajectory which most closely resembles the AU's business objective. This objective is a balance between arriving on time, fuel cost, and departure flexibility. The objective of the Air Navigation Service Provider (ANSP) at the arrival airport is to facilitate that AU's trajectory as much as possible but also facilitate the trajectories of the other AUs. The latter requires optimal use of the available capacity for departing and arriving aircraft. The task of optimising the arrival trajectories of all AUs within the limits of the arrival airport is called *arrival management*.

The capacity constraints at an airport may require adjustments to the trajectories of inbound flights. To fly the optimal trajectory from an AU's perspective, such deviations from these optimal trajectories should be minimal. Furthermore, by the trajectory earlier, the flight efficiency is increased. For example, a more minor speed increase over a longer flight time is often more fuel-efficient than a more considerable speed increase over a shorter time. Yet both speed increases achieve the same result [2].

By planning aircraft arrival times, ANSPs all around the world aim to balance the demand with the available landing capacity [3]. Through timely planning, AUs can fly a trajectory close to the optimal one, given the available capacity. Such timely planning and adjustment of trajectories requires accurate prediction of the ETAs of the flights at the airport and the ability for the ANSPs to adjust the trajectories at those horizons.

Developments in communication technology and specific developments for ATM enable—or are foreseen to enable—influencing aircraft trajectories at much longer time horizons [4]–[6]. This increase in horizon can further increase the potential gains in efficiency from arrival management. The expected arrival time can be predicted by sharing information on flight progress through, for example, Airport Collaborative Decision Making (A-CDM) and System Wide Information Management (SWIM). At the same time, a concept such as SWIM, initiatives such as Single European Sky ATM Research (SESAR), Next Generation Air Traffic Management System (NextGen), and the concepts proposed in the ICAO Global Air Navigation Plan, will allow Air Traffic Controllers (ATCOs) at the destination airport to request trajectory changes of aircraft still under control by upstream ANSPs [6]–[8]. These earlier actions would further reduce the cost of deviating from the planned arrival time by, for example, an even smaller speed change or a decision to delay departure [2].

1.1 Arrival management

Nowadays, many ANSPs use Arrival Managers (AMANs) to support the responsible air traffic controllers—*sequence managers* (Europe) or *traffic management coordinators* (US)—in deciding how to best either modify the 4D trajectory of inbound aircraft or adjust landing capacity [9]. These information and decision support systems provide the sequence manager[1] with predicted arrival times of the inbound aircraft and sometimes support the operator further by automatically planning the optimal arrival times. Figure 1.1 provides an example of the graphical interface of a typical AMAN system.

The accuracy of an ETA depends on three factors in its calculation: The accuracy of the prediction algorithm, the accuracy of the flight model, and the accuracy of the inputs needed to perform the calculation [10]–[12]. Common to all trajectory predictors is that, as the prediction horizon increases, prediction errors accumulate [13]–[17]. Furthermore, a longer prediction horizon may introduce new sources of errors. For example, the exact take-off time becomes a factor in the prediction when the departure at the origin airport comes within the planning horizon.

[1] Depending on the ANSP, the task of sequence manager may be combined with another such as approach controller or flow controller and be designated as such. For clarity, this dissertation will use *sequence manager* for the specific task of planning arrivals.

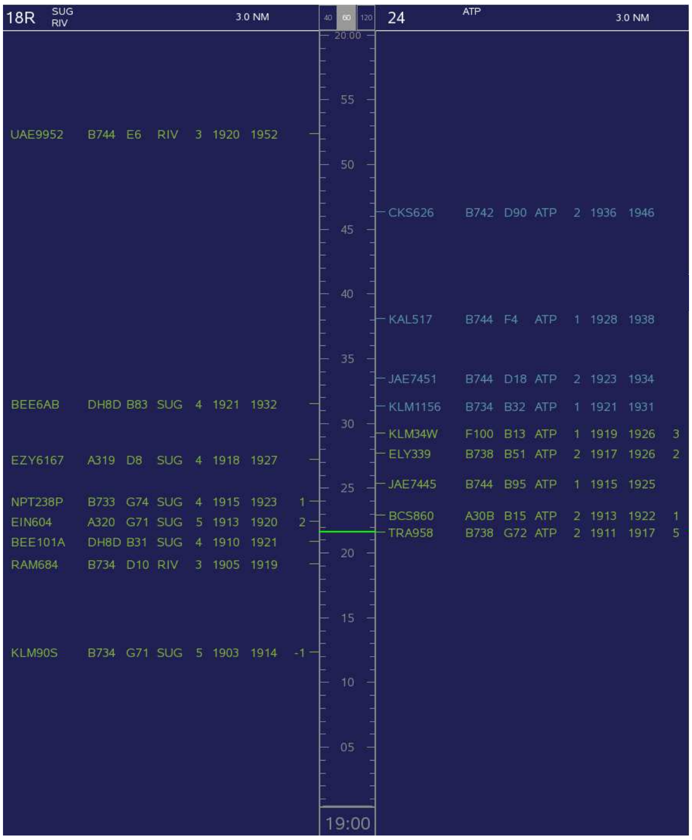


Figure 1.1: Example of an AMAN Human-Machine Interface. This particular display is operational at Air Traffic Control the Netherlands since 2018 (Source: LVNL).

As the uncertainty of the prediction increases, any decisions based on erroneous ETAs carry an increased risk of being incorrect. Such an incorrect decision will require revision of the trajectory later, negating any benefits of early decisions. A planning based on more uncertain information has a higher chance of being non-optimal for the situation at hand. ATC may need to revise the plan to safeguard the balance of demand and capacity, which may lead to further deviations from the planned trajectory. These revisions will further reduce the benefit of arrival management and may even be detrimental to overall flight efficiency. At the same time, the workload for the sequence manager will increase as they have to re-plan the arrival times more often.

At the moment, the standard deviations of prediction error vary strongly and depend strongly on the phase of flight and the source of the prediction. For example, typical predictions by the Flight Management System (FMS) have a standard deviation of 30 seconds at 20 minutes before an arrival point when airborne [18]. On the other hand, the standard deviation can increase to 15 minutes when the aircraft is still on the ground (e.g., the departure accuracies of several airports in the US found by Mueller and Chatterji [19]).

If the planned sequence of aircraft is to remain fixed, the uncertainty of a predicted arrival time has to be well below the landing interval between two aircraft. Typically, a single runway handles between 40 and 50 aircraft per hour [20]. These rates imply that a typical landing slot is slightly longer than one minute. The current highest prediction accuracy of 30 seconds (downlinked from an FMS as introduced above) would limit the horizon to the 20 minutes at which this accuracy is achieved. Those airports at which planning is performed at a longer horizon subsequently control the aircraft to meet that time, mitigating planning problems at the risk of non-optimal trajectories.

AMAN decision horizons in Europe are currently limited to typically 30 to 40 minutes before landing to avoid such non-optimal decisions [3]. At these horizons, the uncertainty in arrival time is small enough to make the likelihood and impact of incorrect decisions acceptably low. This horizon is closely related to the fact that the flight times between many of Europe's larger airports lie just beyond that horizon (see Figure 1.2). Extending the prediction such that flights between those airports would fall entirely within the prediction horizon would drastically increase uncertainty as aircraft 'pop up' in the planning when they have just departed. This indicates that the current uncertainty on the ETAs is too large to support useful arrival management at further horizons. Enabling the improvements in the AMAN process by increasing the horizon, therefore, first of all, requires addressing the problem of prediction uncertainty.

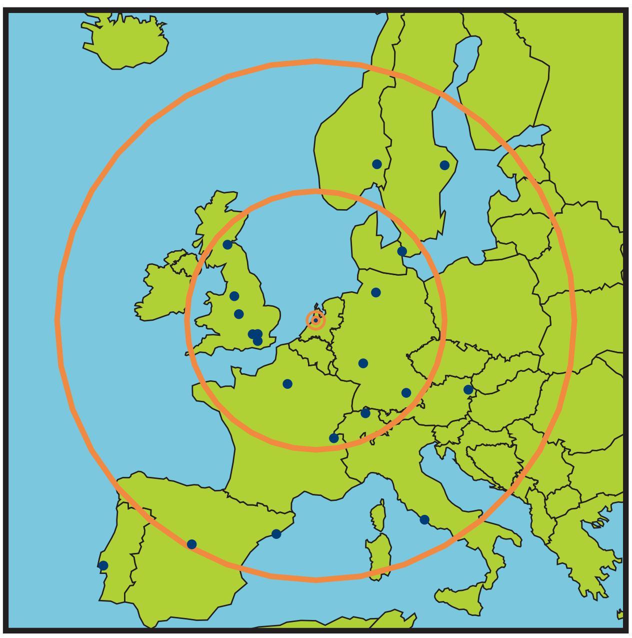


Figure 1.2: Proximity of origin airports in Europe: The map shows the 20 busiest connections into Amsterdam Airport Schiphol in 2013 according to [21]. The range rings indicate 400 and 800 nautical miles or one and two hours flight time, respectively.

1.2 Addressing prediction uncertainty problems

The problem of prediction uncertainty in AMAN has been recognised as one of the limitations to AMAN development for several years [3], [11]. Several projects are therefore looking at solving the problem of arrival time prediction uncertainty and may, therefore, enable a longer prediction horizon [12], [17], [22], [23]. The different solutions to the problem can be divided into two distinct directions: Either increasing the accuracy of the predicted arrival times and therefore reducing uncertainty or, enabling the arrival management process to better cope with the inevitable uncertainty.

1.2.1 Reducing uncertainty

Developments in communication and algorithms will provide more accurate input data to more accurate predictors. The SWIM concept, combined with sharing 4-Dimensional Trajectories (4DTs) of aircraft is expected to significantly reduce the uncertainty on inputs [4], [6], [24]. When this information is furthermore used in new prediction algorithms or predictors that use more accurate models of the flight path, a more accurate ETA can indeed be predicted [11], [25]. Finally, concepts such as Trajectory Based Operation (TBO) and Required Time of Arrival (RTA) make the arrival time a target parameter for the AU. This last step closes the loop: The AU will then try to correct the error in the predicted arrival time, and the uncertainty on the arrival time is only limited by the ability of the AU to meet that time [4], [26], [27].

However, these developments are likely to be limited in their efficacy in the short term for two reasons: First, sharing of more accurate input information depends on whether that information exists or not. Secondly, the implementation of new technology in aviation takes time.

As the prediction horizon grows, more processes will need to be included in the prediction, which introduces more and more potential for errors. An example of such introduced errors are flights that have not yet completed the boarding process. Anyone who has been at a major international airport such as Amsterdam Airport Schiphol will recall the remarkable number of calls for delayed passengers as they are “delaying the flight”. When that happens, the punctuality of the aircraft in leaving the gate depends on the punctuality of the passengers. If a passenger arrives late, the AU must decide to either wait for the passenger or spend time finding and unloading the passenger’s checked luggage. Both choices cause delay.

More strict adherence to the planned trajectory could solve problems such as the one above at the cost of flexibility (e.g., the scheduled boarding time). However, while this would improve accuracy, it would also put further requirements on the operation. In the example of the late passenger, this would mean that the airline would have to spend—and pay for—more time at the gate to allow the passengers to board in time and for the passenger to be at the airport earlier and therefore increase total travel time. At some point, the reduction of flexibility might increase the operating cost or reduce service to the end customer.

The slow adoption of any new technology in aviation follows from both a cautious approach (as any change may impact safety) and the long lifespan of aircraft. The high cost of refitting aircraft with new systems often implies that systems will only change with the replacement of the aircraft. This inertia will both limit the rate at which improvements can be implemented and cause a variety in achievable prediction accuracy during phases of mixed equipage. For example, newer aircraft may be able to share the FMS-predicted trajectory, which has a lower uncertainty. Older aircraft—which will still be flying together with the newer aircraft—cannot provide this trajectory and need to be planned using a less accurate prediction. Since both aircraft are flying simultaneously, the working horizon will be governed by the less accurate prediction.

1.2.2 Coping with uncertainty

Instead of reducing flexibility, the second direction considers the design of an AMAN process that can better cope with uncertainty. This approach intrinsically allows for more time flexibility while still supporting necessary planning. In such a concept, the uncertainty is taken as a given, and techniques aim to allow arrival management while knowing that prediction errors may exist. The goal of this approach is to either plan arrival times so that possible errors can be accommodated easily or use knowledge on uncertainty to decide when a decision is likely to be effective.

In such an approach, the benefits of a longer planning horizon can be achieved by planning whenever uncertainty is low. Conversely, when uncertainty is high, the planning decision may be delayed to avoid having to re-plan at a later stage. Moreover, this method can support the gradual implementation of more accurate prediction capabilities by taking into account the improvement in accuracy when it becomes available.

1.3 Supporting the human operator

Fully automated AMAN concepts are theoretically feasible, and various algorithms have been proposed. However, a literature review of about 20 such algorithms concluded that most have never been implemented in operation [28]. Often, the algorithms were too theoretical and required significant simplification of the operational problem up to a point where implementation would not be useful. Algorithms that would provide sufficient detailed modelling of the operation would subsequently require too much calculation effort to be able to provide real-time information.

At the same time—especially during high-capacity operations—many AMAN decisions are not very complicated. For example, planning aircraft in the predicted arrival sequence often leads to an acceptable solution. Modifications to this solution are only required when sufficient spacing between each landing cannot be guaranteed [28]. In these situations, the necessary planning actions are often routine work. Such routine work can be done by automation.

The above indicates that automated planning algorithms can certainly provide benefits to an AMAN system. However, as the horizon becomes longer, it becomes more and more likely that certain constraints cannot be modelled and introduced in an algorithm. The increase of the prediction horizon and the addition of uncertainty information will make the problem even more complicated. Modelling some factors may never be practically feasible [3]. Think, for example of ad-hoc incidents such as a technical problem with an aircraft departing in front of the flight of interest that delays the departure of that flight of interest.

For these reasons, SESAR foresees a central position for the human ATCO in a future—possibly highly automated—ATM system [4]. These sequence managers will have to make sure that the automation performs as designed and further manipulate the planning using their expertise and knowledge: The operator will not be tasked only to *monitor* the system, but, more importantly, to *manage* it [4, pp. 45].

This role will require the operator to understand the rationale and decisions made by automation and be able to influence these. If they are only provided with the outcome of an automated calculation, the sequence manager may not be able to verify whether the calculation is correct and cannot be considered responsible or accountable for the provided solution. Furthermore, if the operator makes a manual change, the automation may generate further unexpected changes as it is not working along the same strategy as the operator.

If uncertainty is to be included in the decision-making process, a Human-Machine Interface (HMI) will be needed that provides the uncertainty as information to the sequence manager. Secondly, if the human operator is to be responsible, the automation should be a tool for their task. This role requires the automation to communicate its rationale and decisions clearly and understandably.

1.4 Research aim and approach

To enable the benefits provided by increasing the working horizon of AMAN, a means has to be found to work in the presence of prediction uncertainty. The concept should not lead to excessive restrictions on flight operations. Furthermore, the technique should allow a human operator to monitor and control the process. Therefore, the aim of this thesis is:

> To develop (semi-automated) decision support tools (interface and automation) for air traffic controllers in planning arriving aircraft when the predicted arrival times of those aircraft are still uncertain.

To achieve this aim, this thesis has two objectives:

- Develop a method to determine the uncertainty associated with an ETA.
- Design an HMI that allows controllers to perform arrival management in the presence of uncertainty in the ETAs.

These objectives are elaborated in the following approach.

1.4.1 Predicting uncertainty

Prior research on prediction of uncertainty in the ATM domain typically focuses on short horizons (in the order of 20 minutes) for tactical tools or very long horizons for flow management purposes (in the order of several hours) [14], [16], [29]. For AMAN a continuous prediction will be required to estimate the uncertainty from the horizon of two hours until shortly before landing. Because the uncertainty is highly dependent on the horizon, a method is required that takes this variation into account.

Currently, no suitable methods exist to provide an online prediction of ETA uncertainty. To predict the uncertainty for a particular flight, the iFACTS [29] and CARE [16] projects employed methods based on a calculation using knowledge of the uncertainty of the input components. This approach is, however, impracticable in the AMAN context because the large prediction horizon introduces many new sources of uncertainty. Modelling uncertainty requires detailed modelling—and therefore knowledge—of each contributing factor. A method based on empirical data such as proposed by Wanke et al. [30] overcomes this problem but the subsequent Monte-Carlo simulation to calculate uncertainty will require considerable computing time, making the methods less suitable for the real-time function of AMAN.

To predict the uncertainty while avoiding the problems with current uncertainty models requires a model that directly forecasts the prediction uncertainty of a flight based on available information on that flight. Based on estimates currently provided to ANSPs, this thesis proposes an empirical model and validates that model against similar data. This data-driven approach allows for rapid prediction of uncertainty per flight while using a relatively simple algorithm. At the same time, by using currently available information, the technique would be readily applicable in present-day operations and would support a gradually reducing uncertainty as more accurate predictions become available.

1.4.2 Visualizing uncertainty

In a parallel project, interface and automation tools are developed that provide the uncertainty information and enable operators to work with that information. At the moment, some concepts for the visualisation of uncertainty exist [31], but none of these would be suitable for working with an AMAN system.

In the last two decades, and with the help of more advanced computing and visualisation technology, the first operational systems for ATC have been developed that present uncertainty [14], [32]–[34]. These are directed towards tactical separation in particular, working on time horizons that are much smaller than even the present AMAN horizon. Also, these systems focus on *position uncertainty* to support separation rather than *time uncertainty* to support planning.

This research applies techniques based on the Ecological Interface Design (EID) framework developed by Vicente and Rasmussen [35] to visualise the effects of uncertainty on the objectives of arrival management. This information is presented as an addition to the most common form of present-day AMAN displays. Subsequently, the interface and automation are tested in human-in-the-loop experiments.

1.5 Thesis scope

The concept of predicting and working with uncertainty in arrival management is relatively new. The research has therefore been limited to the following initial scope:

- The concept focuses on the intermediate future, i.e., 10-15 years, in which full deployment of concepts such as envisioned in SESAR has not yet been achieved. Uncertainty will be modelled to represent present-day levels of uncertainty.

- The modelling of uncertainty is developed and validated on data from 2013-2014. However, the method is independent of the dataset used and would also support other or more recent datasets.

- The planning problem in this research was initially limited to a single runway or airspace entry point. During the research initial steps have been made to address the problem in case of multiple runways; these will be further addressed in Chapter 5.

- The research will initially focus on the AMAN process as applicable to Amsterdam Airport Schiphol. The complexity of the airspace combined with the nearby location of departure airports makes uncertainty a considerable issue at this airport, as shown in Figure 1.2.

- The proposed HMI is validated in experiments with novice users that are not air traffic controllers as their availability for such experiments is highly limited. Moreover, the operational concept of arrival management at such a large horizon does not exist at the moment and would, therefore, also be unfamiliar to experienced operational staff.

- Since the visualisation is developed in a parallel project to the uncertainty determination, the experiments use a simplified model of the progress of a flight and the associated prediction uncertainty. The relation to actual operation will be discussed in the synthesis in Chapter 6.

1.6 Thesis outline

Figure 1.3 provides a visual outline of the thesis structure. This thesis consists of seven chapters that can be divided into four sections. This and the following chapter set the scene for this analysis. Chapter 2 describes the arrival management operation and work domain. By evaluating the prediction process and the available system support, this chapter describes the relevant gaps in technology when the horizon is to be extended.

The work then splits into two parallel paths: Determination of the uncertainty and visualisation of uncertainty. The two underlying projects have been executed in parallel and independently.

Chapter 3 addresses the first objective of determining the uncertainty in an ETA. This chapter proposes a method based on empirical information and then validates that technique using operational data.

Chapter 4 and 5 describe the development of the interface and its testing in human-in-the-loop experiments. Chapter 4 builds the display using the techniques based on EID for a theoretical single runway. Chapter 5 extends the display to the likely situation for an airport with multiple runways. By making the availability of the second runway a control parameter, the full scope of managing *demand* and *capacity* can be evaluated in the display.

Having developed, and tested an initial concept for working with uncertainty, Chapter 6 will discuss the validity of those findings by bringing the two paths together. This chapter will do so by applying the predicted uncertainty from actual data from Chapter 3 to the proposed interface from Chapter 4. Since the research has been limited in both its time frame of implementation and target airspace, particular attention will be given to the applicability of these findings in the wider future ATM context.

The synthesis is followed by the conclusions of this study in Chapter 7. This chapter will reflect on the previous chapters and draw conclusions and recommendations.

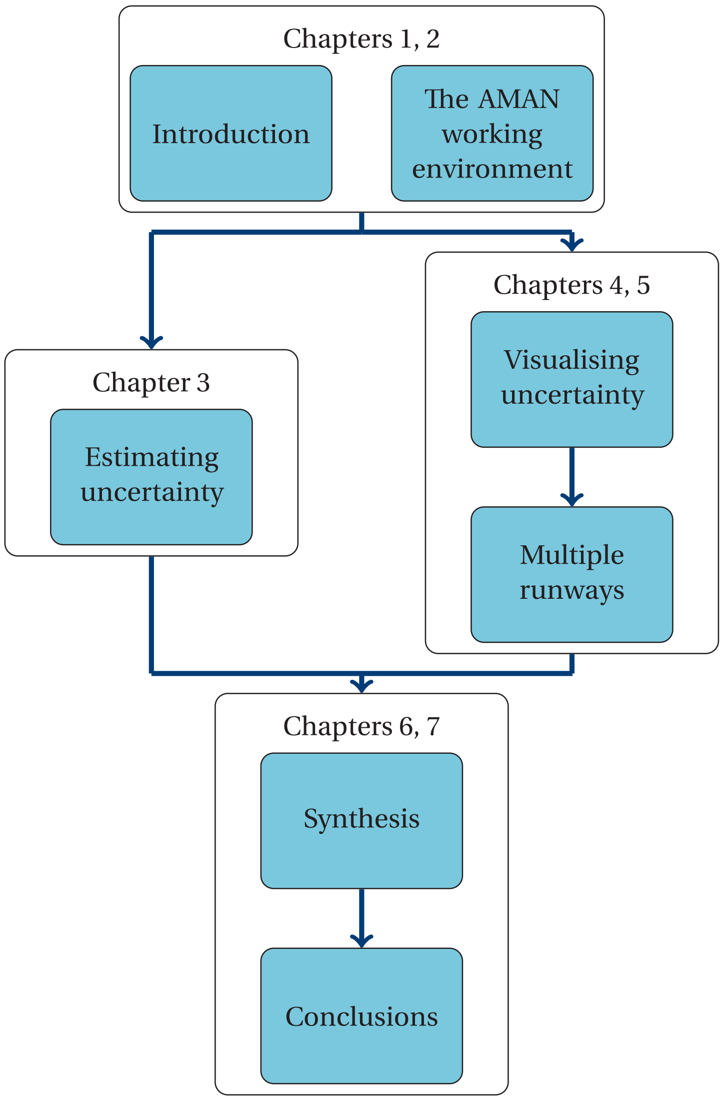


Figure 1.3: Thesis outline

1.7 Bibliography

[1] "Annual Review 2019," IATA, Doha, Qatar, Tech. Rep., 2019.

[2] A. Barff, B. Favennec, P. Conroy, *et al.*, "SESAR P05.06.04 - D28 - Preliminary OSED ed. 00.01.01," SESAR Consortium, Tech. Rep., 2012.

[3] N. Hasevoets and P. Conroy, “Arrival Manager - implementation guidelines and lessons learned,” EUROCONTROL, Brussels, Belgium, Tech. Rep., 2010.

[4] “SESAR concept of operations at a glance ed 02.00.00,” SESAR JU, Tech. Rep., 2011.

[5] JPDO, *NextGen integrated work plan: A functional outline*, 2008. [Online]. Available: `http://www.jpdo.gov` (visited on 03/01/2013).

[6] “Global Air Navigation Plan (GANP) 2016-2030,” ICAO, Montréal, Tech. Rep., 2016. [Online]. Available: `www.icao.int`.

[7] “SESAR Concept of Operations (CONOPS 2019),” SESAR Joint Undertaking, Brussels, Tech. Rep., 2019.

[8] “Concept of Operations for the Next Generation Air Transport System Version 3.2,” NextGen JPDO, Washington, DC, Tech. Rep., 2011.

[9] N. Hasevoets and P. Conroy, “AMAN status review 2010,” EUROCONTROL, Brussels, Belgium, Tech. Rep., 2010.

[10] E. Gilbo and S. Smith, “Probabilistic prediction of aggregate traffic demand using uncertainty in individual flight predictions,” in *AIAA Guidance, Navigation, and Control Conference and Exhibit*, Chicago, IL, USA, 2009, pp. 1–20.

[11] K. T. Mueller, R. Bortins, D. R. Schleicher, D. N. Sweet, and R. A. Coppenbarger, “Effect of uncertainty on En Route Descent Advisor (EDA) predictions,” in *AIAA Aviation Technology, Integration, and Operations Conference (ATIO)*, vol. 6347, Chicago, IL, USA: AIAA, Sep. 2004.

[12] M. Schultz, H. Fricke, T. Kunze, *et al.*, “Uncertainty handling and trajectory synchronization for the automated arrival management,” in *Second SESAR Innovation Days*, Braunschweig, Germany, 2012.

[13] S. Mondoloni, M. Paglione, and S. M. Green, “Trajectory modelling accuracy for ATM decision support tools,” in *International Congress of Aeronautical Sciences (ICAS)*, Toronto, Canada, Sep. 2002.

[14] T. Mueller, J. Sorensen, and G. Couluris, “Strategic aircraft trajectory prediction uncertainty and statistical sector traffic load modeling,” in *AIAA Guidance, Navigation, and Control Conference and Exhibit*, Reston, VA, USA, Aug. 2002.

[15] G. Hunter, “Toward a standardized reference set of trajectory modeling errors,” in *AIAA Modelling and Simulation Technologies Conference and Exhibit*, Providence, RI, USA, Aug. 2004.

[16] D. Nicholls, P. Marti, P. Battino, S. Pozzi, and V. Barsotti, “CARE innovative action presenting uncertainty to controllers & pilots,” EUROCONTROL, Brussels, Belgium, Tech. Rep. December, 2004, pp. 1–53.

[17] M. Paglione, I. Bayraktutar, G. McDonald, and J. Bronsvoort, "Lateral intent error's impact on aircraft prediction," in *8th USA/Europe Air Traffic Management Research and Development Seminar*, Napa, CA, USA, 2009, pp. 527–541.

[18] J. Bronsvoort, G. McDonald, R. Porteous, and E. Gutt, "Study of aircraft derived temporal prediction accuracy using FANS," in *Proceedings of the 13th Air Transport Research Society (ATRS) World Conference 2009*, Abu Dhabi, UAE, 2009.

[19] E. R. Mueller and G. B. Chatterji, "Analysis of aircraft arrival and departure delay characteristics," in *AIAA's Aircraft Technology, Integration, and Operations (ATIO) 2002 Technical Forum*, Los Angeles, CA, USA, Oct. 2002.

[20] T. Morisset and A. Odoni, "Capacity, delay, and schedule reliability at major airports in Europe and the United States," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2214, pp. 85–93, Dec. 2011.

[21] "Traffic Review 2013," Amsterdam Airport Schiphol, Amsterdam, The Netherlands, Tech. Rep., 2013.

[22] R. E. Cole, S. M. Green, M. Jardin, B. E. Schwartz, and S. G. Benjamin, "Wind prediction accuracy for air traffic management decision support tools," in *3rd USA/Europe Air Traffic Management Research and Development Seminar*, Napoli, Italy, 2000.

[23] J. Bronsvoort, G. Mcdonald, M. Paglione, C. M. Young, J. Boucquey, and E. Gallo, "Real-time trajectory predictor calibration through Extended Projected Profile down-link," in *11th USA/Europe Air Traffic Management Research and Development Seminar*, Lissabon, Portugal, 2015.

[24] European Commission, *Commission Implementing Regulation (EU) 2021/116 of 1 February 2021 on the establishment of the Common Project One supporting the implementation of the European Air Traffic Management Master Plan*, 2021.

[25] J. Bronsvoort, G. McDonald, M. Paglione, C. Garcia-Avello, I. Bayraktutar, and C. M. Young, "Impact of missing longitudinal aircraft intent on descent trajectory prediction," in *30th IEEE/AIAA Digital Avionics Systems Conference*, Seattle, WA, USA, Oct. 2011.

[26] J. K. Klooster, K. D. Wichman, and O. F. Bleeker, "4D Trajectory and time-of-arrival control to enable continuous descent arrivals," in *AIAA Guidance, Navigation, and Control Conference*, Honululu, HI, USA, Aug. 2008.

[27] J. K. Klooster and D. de Smedt, "Controlled time-of-arrival spacing analysis," in *9th USA/Europe Air Traffic Management Research and Development Seminar*, Berlin, Germany, Jun. 2011.

[28] M. Mesgarpour, C. N. Potts, and J. A. Bennell, “Models for aircraft landing optimization,” in *Proceedings of the 4th international conference on research in air transportation (ICRAT2010)*, Budapest, Hungary, 2010.

[29] M. Whiteley, “PHARE advanced tools tactical load smoother final report,” EUROCONTROL, Brussels, Belgium, Tech. Rep. November, 1999.

[30] C. Wanke, S. Mulgund, D. Greenbaum, and L. Song, “Modeling traffic prediction uncertainty for traffic management decision support,” in *AIAA Guidance, Navigation, and Control Conference and Exhibit*, Reston, VA, USA: AIAA, Aug. 2004.

[31] R. Finger and A. M. Bisantz, “Utilizing graphical formats to convey uncertainty in a decision making task,” in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Jul. 2000.

[32] P. Whysall, “Future Area Control Tools Support (FACTS),” in *2nd USA/Europe Air Traffic Management Research and Development Seminar*, Orlando, FL, USA, Dec. 1998.

[33] P. Jorna, D. Pavet, M. van Blanken, and I. Pichancourt, “PHARE ground human machine interface (GHMI) project,” EUROCONTROL, Brussels, Belgium, Tech. Rep., 1999.

[34] D. Schaefer, A. Gizdavu, and D. Nicholls, “The display of uncertainty information on the working position,” in *23rd IEEE/AIAA Digital Avionics Systems Conference*, vol. 1, Washington, DC, USA, Oct. 2004.

[35] K. J. Vicente and J. Rasmussen, “Ecological interface design: theoretical foundations,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 22, no. 4, pp. 589–606, 1992.

Chapter 2

Arrival management and trajectory prediction

Finding ways of dealing with uncertainty in arrival management requires understanding the process and its environment. This chapter explores the concept of arrival management, the relevant actors, the constraints on the process, the underlying prediction techniques, and the displays used by sequence managers in current operations. Using that analysis, it will identify the impact of uncertainty and the shortcomings of present-day displays.
This chapter was written at the start of this research between 2011 and 2014. Due to the slow nature of innovation in air traffic control, most of the content is still applicable today. However, some differences may exist. This chapter describes the situation at the time of writing and will explicitly refer to the current situation in 2022 when describing the present-day situation.

2.1 Introduction

2.1.1 Context

In the coming years, Air Traffic Management (ATM) is expected to move from a sector-based, tactical form of control to a more strategic approach based on developing capabilities to accurately predict, share and execute 4-Dimensional Trajectories (4DTs). In this form of control, aircraft will adhere to a strict lateral and vertical route wherever possible, and time will become an explicit control variable. Such trajectories reduce the uncertainty of the future path of the aircraft. The smaller uncertainty subsequently reduces the number of potential interactions between aircraft that may need a resolution. Based on these techniques, it is expected that aircraft will be able to operate close to the Airspace User (AU)'s preference while at the same time minimising delays, environmental burden, and fuel costs due to disruptive short-term tactical intervention [1].

Essential to such 4D operations is the ability to plan arrival times into airports or airspaces with limited capacity. During *arrival management*, different aircraft, which will initially be geographically far apart, will have to be metered, sequenced, and merged onto a limited number of runways. In present-day operations, most adjustments to such inbound flows consist of tactical vectoring and speed adjustments in the final segments of the flight. The resulting lengthening of the flight paths leads to less-than-optimal descents and, therefore, higher fuel consumption. By using predicted trajectories, spacing adjustments may be achieved earlier, possibly even during cruise flight.

Most current concepts assume that an Air Navigation Service Provider (ANSP) performs arrival management for all inbound flights at a central location for each airport or group of airports [2], [3]. At this location, the 4DTs from all aircraft are used to develop, achieve, and monitor an arrival schedule. While fully automated arrival management concepts are theoretically feasible when optimising for some parameters, the complexity of the problem generates considerable difficulty in modelling and controlling all objectives. Furthermore, the uncertainties associated with trajectory prediction (e.g., arrival time) and the uncertainties in the situation (e.g., runway changes or deviations due to weather) may disrupt the—optimal—initial plan. These problems drive a need for a human operator as *automation manager*, not the *automation monitor* [1, pp. 45].

To be able to control, the Air Traffic Controller (ATCO) has to be able to evaluate, compare, and communicate such 4DTs. The controller will need a Decision Support Tool (DST) that provides support in regular operation, in off-nominal conditions, and during recovery to normal operation. In arrival management, the Arrival Manager (AMAN) provides such a DST. Finally, the rollout of such technology and procedures are expected to be gradual. Such a gradual rollout implies that intermediate states will exist in which a mixed operation of Trajectory Based Operation (TBO) and conventional operation needs to be supported.

Especially the last two points show the need for automation that supports the controller in understanding the situation and remaining the key decision-maker in the 4DT operation during the arrival process. Earlier research has demonstrated that all existing systems improve operation when compared to a scenario without DSTs [4]. However, these studies also showed that controllers might not always accept such tools and therefore disregard their advice when they can not comprehend the tool's reasoning [5], [6]. Therefore, this chapter will evaluate the level of support provided in current concepts and will try to identify where improvements are needed in the transition to arrival management based on 4DTs.

2.1.2 Purpose of this chapter

A considerable body of research has been published on potential AMAN procedures and algorithms [7]–[10]. Furthermore, initial reports from Single European Sky ATM Research (SESAR) have established potential operational concepts. However, limited work is available on the human-machine interaction aspects of these procedures and potential Human-Machine Interfaces (HMIs). Finally, information on operational systems in the public domain is mainly limited to commercial publications, which have limited information on the design and use of the interface.

To be able to develop appropriate DSTs, this document evaluates current systems and concepts for future systems. By doing so, it identifies gaps in human decision support. Based on the conclusions of this study, new DST concepts may be investigated to improve support, leading to better operational performance.

2.1.3 Definition of AMAN

No unique definition of AMAN exists [11]. A first definition can be taken from SESAR [1, pp. 50]:

> An ATM tool that determines the optimal arrival sequence times at the aerodrome or possibly at other common route fixes (e.g., IAF).

This definition focuses on the purpose of AMAN as a sequence planner. This definition does not address, and possibly excludes, the role of a human controller. As will be demonstrated later on, such a statement might conflict with the envisaged central role of the human operator as *manager* and not *monitor* of the ATM automation as envisaged by SESAR.

In a survey on existing AMAN systems by EUROCONTROL, the definition used states the role as DST but does not state a clear purpose [12, pp. 14]:

> …to provide electronic assistance in the management of the flow of arriving traffic in a particular airspace, to particular points, such as runway thresholds or metering points.

This study will use a combination of the above definitions to define both purposes to achieve an optimal flow and sequence, and to define the system as a DST:

> A system that provides electronic assistance to a human decision-maker in developing and monitoring an optimal arrival schedule at the aerodrome and/or other common route fixes in the route toward the aerodrome.

2.1.4 Scope

This project aims to investigate concepts for AMAN systems in de medium to long term, between 2025 and 2035. This scope includes the following operational concepts:

- Current operations based on tactical, vector-based control, supported by advanced AMAN,
- Intermediate operations in which a subset of the participating airspace users will be controlled using 4DTs as envisioned in SESAR,
- Full TBOs in which AMAN is still governed by ground-based Air Traffic Control (ATC).

In all cases, it is assumed that human ATCOs will have the authority and remain in control of the planning process, as explained at the start of the chapter. The study focuses on the automation and the HMI given the above types of operation. While the procedures change, this study assumes that the division of responsibilities remains the same: A *sequence manager*[1] will have the responsibility to plan the arrivals of the inbound aircraft and provide that schedule to executive controllers who issue the instructions or trajectories to the flight crew.

Since this chapter is part of a project supported by the Dutch aviation sector, the primary focus will be on the Dutch airspace and similar airspaces.

[1] As stated in Chapter 1: Depending on the ANSP, the task of sequence manager may be combined with another such as approach controller or flow controller and be designated as such. For clarity, this dissertation will use *sequence manager* for the specific task of planning arrivals.

2.1.5 Document structure

Section 2.2 explores the operational problem by looking at the environment and the constraints that drive the need for an arrival planning process. In Section 2.3 we will explore the arrival management process. Sections 2.4 until 2.6 the describe the underlying systems: Trajectory Predictors (TPs), the HMIs and the possibilities for automation. Using the knowledge on the actual AMAN operation and the automation that supports it, Section 2.7 evaluates the DSTs for their suitability in the current operation. Section 2.8 then concludes this evaluation and describes several areas of improvement.

2.2 Working environment

Near airports, aircraft from all directions will have to be metered, sequenced and merged to a limited number of runways. During cruise flight, the primary purpose of ATM is to ensure safety by keeping aircraft separated while maintaining an efficient flow. During approach—especially during peak capacity scenarios—this objective conflicts with the need to bring aircraft closer to each other to maximise throughput. An arrival management process is needed when the available capacity can (temporarily) not accommodate the offered flow (i.e., two or more aircraft are predicted to arrive at the same time). Therefore, AMAN first and foremost aims to balance capacity and demand.

As aircraft approach their destination, the following objectives will need to be met:

- The aircraft has to descend, decelerate, and be prepared for landing,
- The aircraft has to be scheduled and assigned to a runway,
- The aircraft has to be metered to arrive according to the schedule,
- The aircraft has to be merged into the flow of aircraft to that runway,
- The aircraft in the flow has to be spaced from the aircraft ahead and behind,
- The spacing has to be monitored and maintained until landing,
- The aircraft has to land, slow down and vacate the runway.

This section explores the physical world in which this process takes place. This environment includes the aircraft, the airports and the airspaces but also the human operators acting in that environment. The section then analyses the different capacity limits.

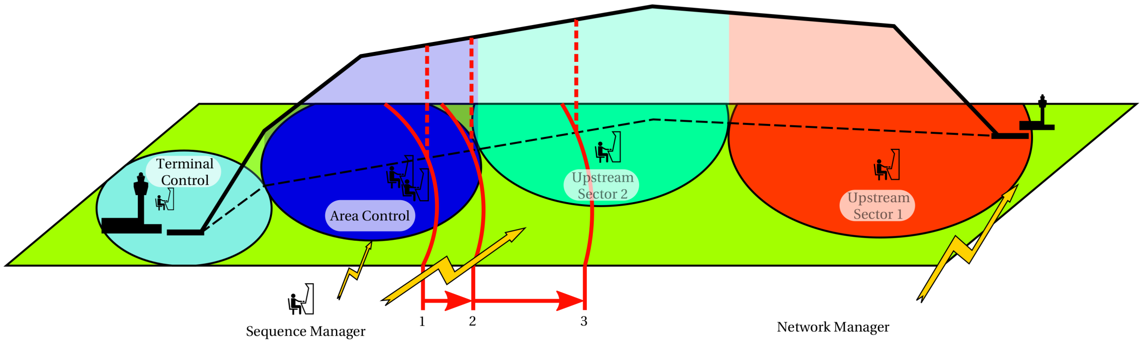


Figure 2.1: Diagram showing the role of the sequence manager: providing planning information to the executive arrival controllers. As the AMAN horizon expands, the planning information will need to be passed to more controllers (1: typical current situation, 2: reaching the end of the ANSP's area, 3: reaching beyond the ANSP's area into upstream sectors).

2.2.1 The arrival airspace

Before describing the arrival management process, it is important to understand the environment in which the process acts. Therefore, this section will describe the objects and the actors in the process. The actors are depicted in Figure 2.1.

Airspace users

The primary customers of the ATM system are the aircraft that use the airspace. For AMAN, the key AUs are those aircraft that approach a particular airport or group of airports. An optimal trajectory for the aircraft is that trajectory which is closest to the operator's plan. At the same time, the dynamics and limitations of the aircraft constrain the available trajectories during the arrival process. During the approach, the operator's intent is a balance between the costs of fuel versus the costs of deviating from the arrival time that the AU had scheduled for that flight (i.e., the Scheduled Time of Arrival or STA).

While fuel costs depend on the profile's efficiency, the cost of deviating from the Scheduled Time of Arrival (STA) depends on the operator and even on the individual flight. The balance, and therefore the optimal profile, can differ per flight. Even more, this balance may shift during the flight if the flight starts deviating from the original planned trajectory. A good example would be the preferred profile once the flight is delayed: An operator with several transit passengers may prefer to spend more fuel to minimise the delay. When the delay does not affect profitability, an operator may choose to adhere to the original profile and arrive later but save on fuel costs.

While ATC often provides flights with an initial route and descent, the optimal descent from the AU's perspective is best planned from cruise to landing. To fly an optimal descent, the crew has to program the aircraft's Flight Management System

(FMS) before Top of Descent (TOD). As the preparation includes the selection of route and runway, these should ideally be provided before TOD. The aerodynamic efficiency of modern airliners limits the maximum descent angle if an aircraft is to maintain or reduce its speed. Therefore, a typical descent has to start between 70 and 120 NM from the destination, depending on the aircraft type [13]. At ordinary speeds, this leads to a time horizon of 20 to 30 minutes. To support the AUs in their optimisation, the AMAN system should support planning from at least that distance and time from the destination airport.

Airports

The focal point of an arrival management process is the airport. All aircraft will have to land on a limited number of runways, slow down, and taxi over a limited number of taxiways to a limited number of stands. The layout of the airport's manoeuvring area (the runways and taxiways), weather conditions, and the local surroundings limit capacity.

In most cases, the airport operator is a different entity than the ANSP. Ultimately, the operator decides how many aircraft it is willing to accept and which runways to make available. During operation, the airport operator may also need to close the runway for maintenance (at longer notice), inspection (at short notice), or for removal of ice and snow (short notice). At all times, the operator will do so in communication with the ANSP.

Aircraft preferably take-off and land into the wind. While an airport may have multiple runways, the wind affects which runways that are useable from an operational point of view. Similarly, AUs have requirements on the length and equipage of a runway in accepting it as a landing runway.

Finally, the environmental effects of aviation may lead to local restrictions on operations. At many airports, the regulator has limited operations to reduce the adverse impact of aviation on the local population. Limits can be applied to the number of flights, the use of runways, or the use of approach routes. Examples are the night curfew at Frankfurt, the night restrictions and daily runway switch at Heathrow, and the preferential runway system and mandatory night time Area Navigation (RNAV) routes at Amsterdam [14].

Airspaces

As described earlier, a typical current arrival process starts at least 70 to 120 NM away from the airport. This distance means the aircraft will traverse a large distance and cross several different airspaces. Especially in Europe, the typical AMAN time horizon is beyond the location of many nearby airports. Therefore, the aircraft in the AMAN schedule may not even have taken off while inside the horizon (See Figure 1.2).

Figure 2.2 gives an example of the traversed airspace structure. The figure shows different types of airspaces, each of which has its own role and responsibilities. For the arrival management process, the following airspaces are of importance, going from the last to the first:

- The control zone, or tower airspace (TWR), controls the airport, indicates what capacity the airport can handle, which runways are available, and participates in deciding which runways to use.
- Approach Control (APP) guides the aircraft to their designated runways and spaces them for optimal use of the available capacity. In this Terminal Maneuvering Area (TMA), aircraft will—inherently—come closer together. The aircraft will have to turn from a great circle between origin and destination to the runway direction. Finally, the approaching and descending aircraft have to share their airspace with departing climbing aircraft and aircraft transiting the airspace. The complex nature of this problem often makes approach control the deciding party on the desired sequence and schedule of aircraft.
- Area Control (ACC) descends and guides the aircraft to the approach control airspace and aims to deliver the desired sequence and already close to the desired schedule.
- In the Upper Area Control (UAC), aircraft are mostly in cruise flight. However, as the AMAN horizon increases, the UAC may already be involved in providing information to the AU and working toward the planned schedule.

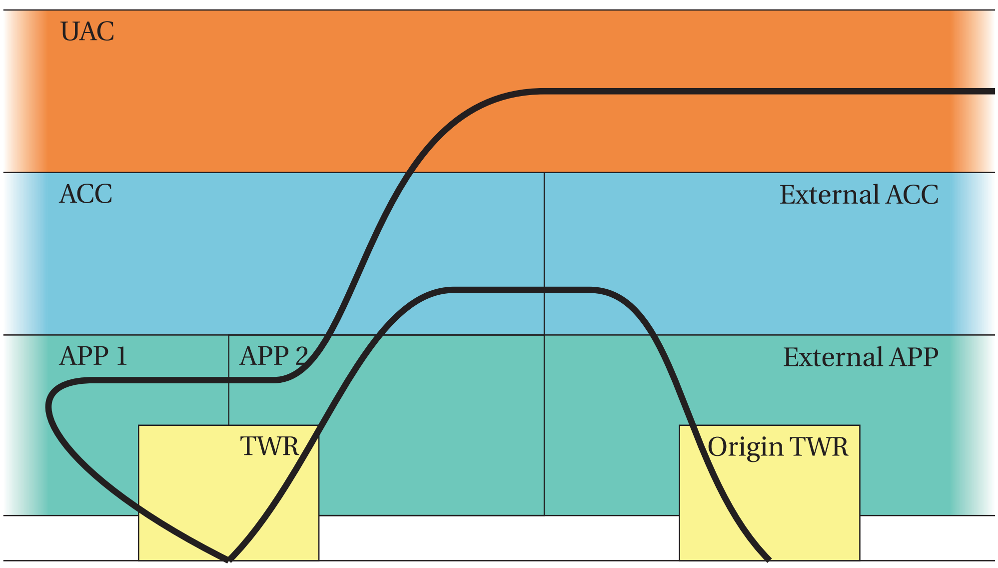


Figure 2.2: Cross section of airspaces for two typical flights to Amsterdam Airport Schiphol showing how different flights may be crossing different sectors during their approach.

Note that the different airspaces may well fall under the responsibility of different ANSPs and may be in separate sovereign states. Therefore, the arrival management process will require communication within the ANSP but often also between different organisations.

As some airspaces may be reserved for specific purposes (e.g., military use, air sports, departing traffic), the boundaries may form temporary or permanent constraints on the possible paths of aircraft. Finally, certain weather phenomena may lead to temporary obstacles to the path of aircraft.

2.2.2 Constraints

As described in the introduction of this chapter, the arrival management process is required when the demanded capacity is more than the available capacity, even if only for two aircraft. The environment described in the previous section introduces constraints on the available capacity to receive arriving aircraft. This section will describe those constraints and their horizon. The horizon is the time between the determination of the constraint and the landing of the first affected flight.

Airport capacity

The key bottleneck in arrival procedures is the destination airport (See Section 2.2.1). This infrastructure defines a hard limit in the number and direction of runways available. Furthermore, the taxiway infrastructure limits aircraft movement rate on and off the runway. For example, some airports do not have a taxiway parallel to the runway, requiring the landing aircraft to turn and backtrack to the runway exit before the next aircraft can land.

While long-term maintenance follows from a plan, runway closure is also a function of weather conditions (wind, snow) which may change within the AMAN horizon. Secondly, the use of the runway may be governed by regulations to limit local environmental effects below the flight path to that runway (e.g., at Amsterdam [15] or Heathrow [14]).

Note that runways often are selected as a function of required capacity. Runway planning, therefore, becomes a part of the AMAN process and has to be performed in collaboration with the airport operator.

Separation

The primary purpose of ATM is to prevent collisions between aircraft while maintaining an efficient flow, both in the air and on the manoeuvring area (i.e., the airport surface) [16]. Several types of separation limit the minimum time between two landing aircraft:

- Runway separation: No two aircraft may simultaneously be on a single runway.
- Radar separation: Aircraft should be sufficiently far apart for ATC to detect a potential collision, to communicate avoiding action, and for the aircraft to execute that action.
- The third type of separation does not relate directly to actual collisions but to the effect of wake turbulence: Flying aircraft leave trailing vortices; following aircraft have to maintain a minimum distance to avoid control problems.

Runway separation

An aircraft is only allowed to land on—or take off from—a runway when no other aircraft or vehicles are on the runway. If a runway is used for arrivals and departures simultaneously, an aircraft is only allowed to land when the departing aircraft has crossed the runway threshold or commenced a turn. For airports with multiple runways, the runways or their approach and departure routes may intersect. In these cases, separation should also be assured between aircraft on different runways [17].

The time between each landing has to be sufficient for the leading aircraft to land, slow down to a safe taxi speed, and leave the runway before the next aircraft is allowed to land. When runways are used for departures as well, the runway occupancy is determined by the time the departing aircraft needs to line up, perform the final preparations for take-off, accelerate to take-off speed, and take off. In either case, the speed at which aircraft can enter or leave the runway depends on taxiway layout, aircraft type, aircraft equipage, and operating procedures.

Low visibility conditions may require local ATC to take extra time to ascertain that aircraft have vacated the runway. Such conditions may also reduce the acceptable combinations of runways as aircraft separation can no longer be monitored visually.

While possibly different per flight, runway capacity is mainly a function of the selected runway and the weather conditions. These conditions typically change in 15 to 30-minute time scales.

Radar separation

A minimal distance between aircraft is applied to support ATC in preventing collisions in the air. This distance is based on the ability to determine the aircraft's position using surveillance equipment and the time required to detect and avoid a potential collision by ATC. In modern TMAs this is typically 3 NM horizontally (based on radar accuracy and precision) and 1000 feet vertically (based on the accuracy and precision of barometric altitude measurements). In en-route airspace, further from the radar heads, the required separation is increased to 5 NM horizontally. Finally, 2.5 NM may be applied at some airports during final approach [18], [19]. The air safety authorities set these standards, and they will not change the standards during the planning process.

Wake turbulence separation

Within a stream of arriving aircraft, the effects of turbulence generated by the aircraft further limits minimum separation. The strength of the turbulence depends on various factors, most notably the weight of the aircraft, and diminishes with time. The sensitivity to the turbulence of the following aircraft depends on multiple factors, with weight being the most notable again. Therefore the required separation—and thus the landing rate—is governed by Wake Turbulence Category (WTC) of the pairs of consecutive aircraft, as shown in Table 2.1. Note that these are the separation standards as described by International Civil Aviation Organization (ICAO), some states have opted for slightly modified categories or distances (e.g., UK [17], [20]).

Table 2.1: ICAO separation minima.

WTC leader	WTC trailer		
	M	H	J
M	rad	6 NM	7 NM
H	rad	4 NM	6 NM
J	rad	4 NM	4 NM

M = Medium, H = Heavy, J = Super
rad = radar separation minimum, (typically 3 or 2.5 NM) [17], [21]

The effect of these constraints on capacity depends on the selected sequence of aircraft. The 4th, 5th and 6th row of Table 2.2 demonstrate this effect on throughput; all three sequences contain an identical fraction of *heavy* and *medium* aircraft. The sequence of alternating classes has the lowest throughput rate. The most grouped scenario (one series of mediums, followed by an equal number of heavies) has the highest rate.

As the required separation between aircraft is currently expressed in distance, the time between each aircraft also depends on the ground speed of the trailing aircraft. The ground speed depends on the aircraft itself and on the headwind component; An increase in headwind reduces the landing rate, as shown in Table 2.2. Therefore, the effect of the standard depends on the sequence which is controlled by the AMAN process, and the headwind component which may change in 15-30 minutes (see Table 2.2).

Table 2.2: Throughput as function of the arrival sequence.

Aircraft sequence	headwindcomponent [kts] 0	10	30
	[aircraft per hour]		
MMM …	55	52	45
HHH …	42	39	34
JJJ …	42	39	34
MHMH …	37	35	31
MMHHMMHH …	44	40	35
MMM …HHH …	46	43	39
MJMJ…	33	31	29

M = Medium, H = Heavy, J = Super
Assuming a speed of 160 knots which is typical at about 4 NM from the runway.
ICAO separation minima with minimum radar separation of 3 NM [17], [21].

Two developments aim to reduce the constraining effects of wake turbulence separation [22], [23]: First of all, RECAT-EU and Pair-Wise Separation increase the number of wake turbulence classes to reduce pairwise separation when possible. Secondly, Time-Based Separation compensates for the reduction in landing rate due to headwind and flight speed.

Regulations

This section will describe the capacity limits due to regulations imposed to limit the effect of aviation on the local environment. Such regulations can include [14]:

- Curfews, defining periods in a day in which aircraft, or particular types of aircraft, are not allowed to land at all. Such a curfew is commonly applied to limit noise during the night.
- Preference (such as in Amsterdam) or enforcement (such as in Heathrow) of certain runway combinations [24], [25],
- Limits on the routes through the TMA,
- Mandatory use of Continuous Descent Operations (CDO);

Curfews determine whether an aircraft is allowed to arrive at all and are not relevant for AMAN. Runway preferences will impose requirements, as the selection of the appropriate runway for a particular flight will be influenced by the preference, but also by the number of aircraft already planned to that runway.

Routing limitations and CDOs may reduce the number of options for ATC to influence spacing after a certain point in the approach. The reduction in control space will increase the required margin between aircraft, which reduces landing capacity [26], [27].

Controller task demand

The capacity limits described in the previous sections all involve known quantities or rules. A final set of limitations follows from the need to maintain an acceptable workload for the controllers that need to handle the aircraft. Note that this excludes the workload of the operator of an AMAN system. Several attributes may affect the workload [28]:

- The number of aircraft: For example, a higher runway capacity often requires more intervention by the controllers. Therefore, workload typically increases with an increasing number of flights.
- The routes for these aircraft: Intersecting routes and merging of different streams require more attention on separation.
- The presentation of the traffic: Traffic coming from a single direction will already have a natural sequence and will not need much merging.
- Stability of the whole situation: For example, the presence and movement of convective weather will require more attention and action by the ATCOs to support aircraft in avoiding weather.
- System support: Failure of support systems may increase the task demand load at similar traffic levels.
- Team competency, which determines the acceptable task load: This may depend on experience, fatigue due to a long period of high task load, and equipment availability.

The list shows attributes which may be determined objectively (e.g., the number of aircraft, the routes, the available staff). However, some of these criteria may require online subjective judgement by expert supervisors (e.g., traffic presentation, team competency).

2.3 AMAN Process

Section 2.2.1 described the environment in which the arrival management process takes place. This section describes the objectives of the AMAN process and the different steps in the process.

2.3.1 Objectives

The potential benefits of AMAN to the AU can be described using several key performance indicators for SESAR [29]:

- **Safety:** The amount of risk of air traffic accidents. During the arrival process, safety imposes some limits on the environment by introducing separation requirements, as explained in Section 2.2.2. Additionally, safety is affected by the effect of the inbound flow on the controller workload, as a high workload is considered a safety risk. AMAN enables planning for the required separation and may support workload management (see Cost-Effectiveness).

- **Capacity:** The number of aircraft that can execute their flight per unit time. The introduction of this chapter already stated that the principal purpose of an AMAN system is to plan the arrival schedule such that the capacity meets the demand.

- **Efficiency:** The costs of deviating from the Shared Business Trajectory (SBT) as planned by the AU in both fuel as well as arrival time. Improved planning can reduce operating cost to the AU. Following Section 2.2.1, this can be translated as allowing the flight's execution to be as close as possible to its preferred trajectory. Note that this definition of efficiency only encompasses the trajectory of the individual AU; it does not include the cost of ATC.

- **Cost-Effectiveness:** The cost of ATM service to each flight. This cost is, amongst others, related to the number of ATCOs needed to handle the traffic safely (see Section 2.2.2). A higher workload requires more, or better trained, controllers, which translates into a higher cost to the AU. AMAN can provide the support to adjust the flow of aircraft—both in number and traffic presentation—to reduce the workload of downstream controllers.

- **Environmental sustainability:** Enabling aircraft to fly CDOs reduces the effect of flights on the local population. Note that this objective is related strongly to the AU's objective of minimising fuel consumption. Furthermore, due to the lack of control space during the procedure, the ability to provide CDOs—with associated noise benefits—depends strongly on initial spacing, which can be improved through AMAN [27], [30].

- **Predictability:** This parameter measures the degree to which the planned capacity can be made available despite disturbances. For example, weather or sudden runway closures may disrupt a planned schedule. AMAN may help in reducing the effect of such disturbances and shorten the time needed to return to normal operations lowering the cost of disturbances to AUs.

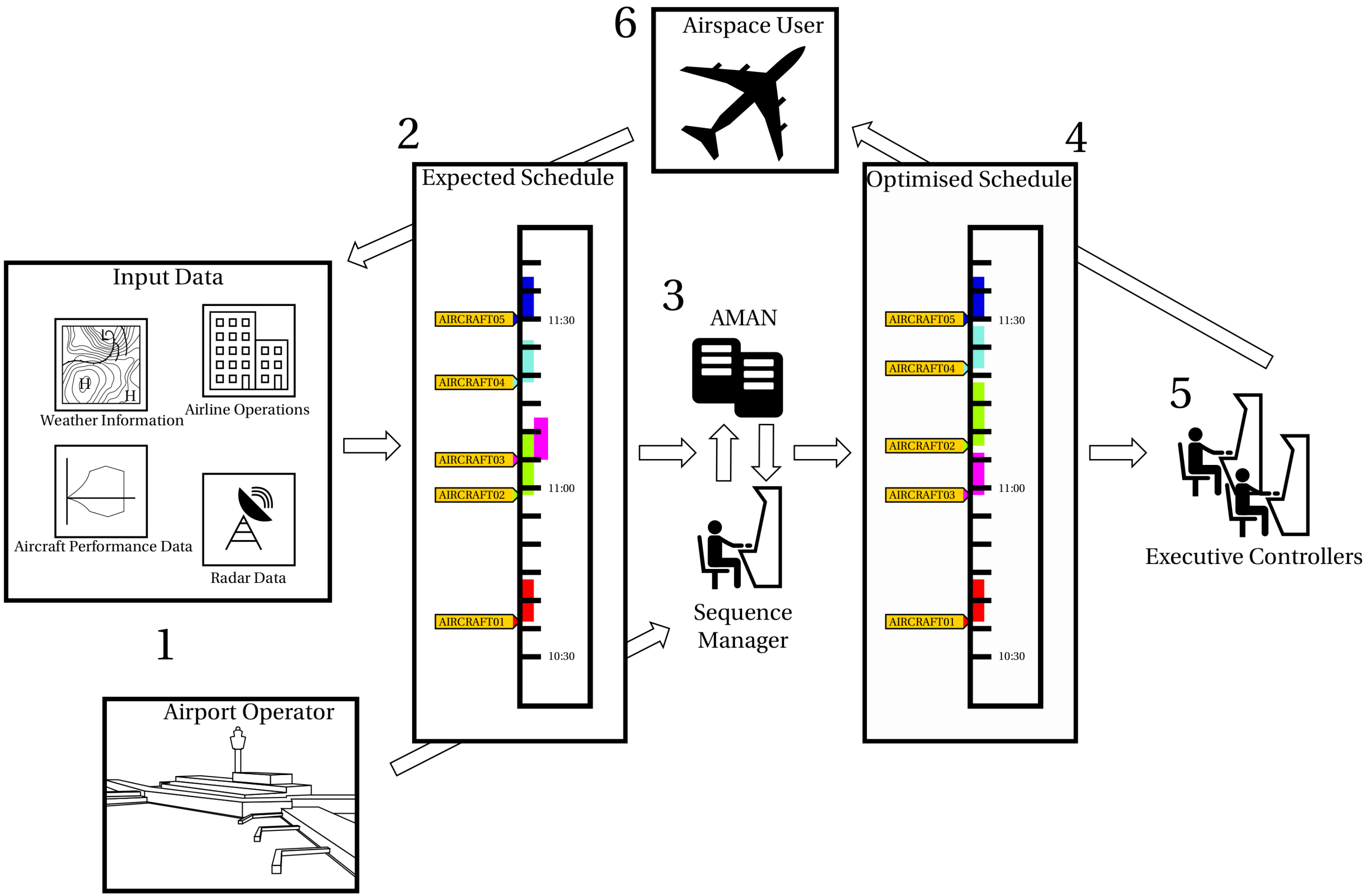


Figure 2.3: Overview of the AMAN process and its actors. (1) Based on available data on a flight, the system calculates an expected schedule (2). The expected schedule and runway availability from the airport operator form the inputs for the sequence manager. They then modify the schedule using the AMAN system (3). The resulting planned schedule (4) is communicated to the executive ATCOs (5), who will communicate with the AU to realise the plan (6).

2.3.2 Actors

The process is a collaboration between several actors (see Figure 2.3):

- The AU, who will have a preferred trajectory and ultimately executes the instructions given by ATC,
- The airport operator, who will have an operational plan for the airport,
- In most operational systems, the sequence manager performs runway planning and scheduling. Depending on the facility, this may be a dedicated function (for example, the Traffic Management Coordinator (TMC) in the US [31]) or a task performed by a supervising role (the TMA supervisor at Amsterdam Airport Schiphol). As this operator develops the arrival schedule, most of the objectives will be relevant to the sequence manager.
- The executive controllers who will realise the schedule by providing instructions to the flights. These are further divided in upstream controllers and downstream controllers, where the downstream controllers are those whose workload is managed through AMAN.

2.3.3 Planning horizons

The arrival management process is commonly divided into several horizons [11], [13], [32]. While the different concepts may have more horizons, the following applies to all concepts:

- The distance or time at which an AMAN system receives information on a particular flight. At this point, sufficient is known about a flight to predict its arrival time. This point depends on the availability of information and the communication of that information to the AMAN system. An example of the current prediction horizon for most systems is the radar horizon [11], [33].
- The distance or time at which the Estimated Time of Arrival (ETA) of a flight is suitable for use in the AMAN process. The location of this horizon depends on the reliability of the ETA but also of the knowledge of the constraints at the destination (e.g., the available runways).
- The distance or time at which ATC can influence a flight to realise the schedule. This distance depends on jurisdiction (the responsible ANSP) and the means to communicate the schedule to the responsible controllers or aircraft.
- The distance at which modification of the schedule does no longer provide a meaningful benefit. At this distance, possible changes to the flight path are either too short (making the change irrelevant) or too drastic (making the change inefficient).

- The distance at which the flight can no longer be influenced. This happens when the aircraft passes the metering point to which the aircraft are planned.

2.3.4 Runway planning

The arrival process starts with planning the runways that will be available for arrival and departure. The responsibility for this decision is shared between the airport operator and the ANSP (see Section 2.2.1). The decision is based on the available runways and the required capacity. Since the runway plan requires information on demand, it is only possible to make a runway plan once sufficient information on inbound and outbound flights is available.

Typically, runway planning is completed at least 30 minutes before the arrival of the first aircraft for that new plan. At this phase, the precise ETA is not yet available but is typically known to an accuracy in the order of 5-10 minutes, providing a rough overview of traffic demand. At the same time, the runway planning strongly depends on the weather, which may also limit the practical runway planning horizon to 30 minutes (see Section 2.2.2).

During the switch from one runway plan to another, arriving aircraft planned to the old combination are in the TMA together with aircraft planned to the new combination. This mix implies that multiple, possibly conflicting route structures are operational at the same time. The increased complexity leads to a higher task demand load and needs to be prepared by the controllers. Often, ATC will use or generate a small gap in traffic to limit problems during the transition phase.

2.3.5 Sequencing and scheduling

Once the runway plan is available, the inbound aircraft are assigned to runways and sequenced based on their ETAs. This phase forms the core part of the arrival management process. While sequencing (i.e., deciding on the sequence of aircraft) and scheduling (i.e., deciding on the individual Planned Times of Arrival, PTA) may be recognised as different processes, their execution is mostly an integrated process.

Assigning an aircraft to a particular runway may be driven by:

- Local procedures. For example, at Amsterdam, where flights from the south and west are normally land on the western runway while aircraft from the east land on the eastern runway. This division reduces the complexity of the routes inside the TMA and thus task demand.

- Wake turbulence segregation to different runways, thus eliminating the larger spacing intervals, as shown in Table 2.1. For example, trials at Heathrow saw the A380 land on the departure runway to limit the impact of the larger spacing requirements on landing capacity [25].

- Runway preferences based on airport layout, aircraft type, and operator preference.

Note that the delays required to achieve a better sequence may void that sequence's advantage. For example, a *medium* following a *heavy* requires 90 seconds separation at 160 knots whereas the reciprocal requires 68 seconds. The advantage of switching the two aircraft may thus be 22 seconds of spacing interval. However, if the *medium* is initially predicted to arrive more than 22 seconds after the *heavy*, no net benefit is achieved as the total delay does not decrease.

Nevertheless, even if total throughput would increase, individual AUs should not be punished unfairly for having a different WTC than other aircraft arriving at that time. An example would be a single heavy in a group of mediums. The—mathematically—optimal sequence would see the heavy as the last aircraft to land (see Table 2.2), requiring it to be delayed until the last *medium* has overtaken it. Similarly, not all sequence switches may be available as this may require aircraft to overtake each other while on the same route, thus risking a separation conflict in the intermediate phase.

Finally, it is worthwhile to note that scheduling does not necessarily imply setting fixed arrival times. For example, when all aircraft that are scheduled to a single runway come from the same upstream sector, the absolute times may not be relevant. Often the upstream sector will build a string of aircraft in which only relative separation is monitored. These "trains" reduce task demand for the upstream controller, who can focus on separation alone instead of separation and arrival time. At the same time, it also reduces the task demand for the downstream controllers as the aircraft have already merged.

2.3.6 Metering

Once a schedule is established, the upstream controllers can start metering the aircraft to the desired times or spacing. The available techniques to adjust the ETA or separation is out of the scope of this study. However, it is worth noting several general properties that affect the control space available to influence the aircraft.

- There is no physical limit to the amount of time that a single flight can be delayed except for the flight's endurance. Delay beyond the endurance limit is also achievable in the form of diversion to alternate airports. Of course, this disregards the AU's objectives.

- The above statement does not hold in a situation with multiple aircraft. At some point, the airspace will be full, meaning that any further addition of aircraft will lead to an unacceptable chance of loss of separation.
- To make an aircraft arrive earlier than predicted (*time advance* or *frontloading*), the aircraft has to increase speed, or it has to fly a shorter route. The first solution is limited by the maximum safe speeds of the aircraft, the latter by the availability of shortcuts in the route.

2.3.7 Sidenote: the metering point

When performing runway planning, the runway is the relevant point of interest for planning traffic. Maximum capacity is achieved when aircraft land at that sequence's minimum total spacing interval. For metering, however, the point for which ATC plans a schedule is important and influences the scheduling strategy.

The choice of the metering point is driven by the need for ATC to control traffic after the planned arrival time is frozen: From a flight efficiency objective, planning to the runway threshold gives the most options for flight optimisation. From a capacity objective, it may be better to meter to an initial point and then relax the tight constraint in time in exchange for more control in spacing.

During the approach, the relative importance of different objectives of arrival management may change. In the planning stage, most aircraft are geographically far apart, and the instantaneous separation is not likely to be a problem. The other objectives drive the arrival process, with predicted separation as a parameter. During the final stages of the approach, the distance between aircraft will be approaching the minimum intervals, and separation becomes a high-priority issue. Effectively, the AMAN process transitions from scheduling to spacing during the approach [34].

Scheduling and metering to the runway threshold enables the precise definition of the arrival times based on the required separation between the aircraft. Once ATC has set a flight to arrive at a planned time, it tries not to interfere with the planned path. The AU can perform an optimal approach to the runway, given the constraints. However, any uncertainty in the predicted trajectory will pose uncertainty in separation during the approach. Therefore, an additional separation buffer has to be applied to compensate for these uncertainties. Such a buffer reduces capacity considerably [26], [35].

A higher throughput may be achieved by influencing traffic until close to landing. Such actions will, however, reduce the efficiency of the flight. In many operational systems, metering is done up to the Initial Approach Fix (IAF), after which approach controllers perform final corrections. Simulations have shown that a combination of the above may work best for high-capacity operations: the acceptable flow should be set using the landing rate at the runway. Subsequently, this rate can be delivered to the upstream metering fixes using an approximation of the flow rate [36].

2.3.8 Merging and Spacing

Once aircraft enter the TMA, approach control has to merge the aircraft onto a single track to the runway and establish and maintain sufficient spacing. While the work itself is outside the scope of this study, the delivered schedule may strongly impact the ATCO's ability to achieve this and thus impact task demand. Examples of properties that affect task demand are the number of merges required, the initial spacing of the aircraft, and the amount of aircraft.

2.3.9 Future concepts

The technological advances in computing, flight guidance, and communication of the last 30 years are expected to enable new forms of air traffic control. This section will briefly examine concepts that may change the arrival management process described in the previous sections. These changes will change the available information, the way flights may be planned into an arrival schedule, and the way the flights may be controlled to realise that schedule.

Information

System Wide Information Management (SWIM) will enable continuous sharing of all relevant information concerning a flight between all involved actors [1], [3]. Implementation of SWIM would make the prediction horizon only dependent on the existence of information but no longer on whether the information is collected by the ANSP that operates the AMAN.

Secondly, AUs will share their planned 4DT both before and during the flight. This information can provide a better insight into the plans, the ETA, and the preferences of the AU [1], [37]. Through sharing trajectory information, the accuracy of the ETA can improve, which could allow for extending the horizon at which information is sufficiently reliable for planning. The next chapter will further explore the effects of uncertainty.

Planning

Currently, the departure time of a flight in Europe is governed by the EUROCONTROL Network Manager (NM), which verifies whether the flight will not exceed capacity on any downstream sectors or the arrival airport. If the flight crosses a capacity-constrained airspace, it receives a window—of currently 20 minutes—in which it has to depart [38]. If the departure airport is within the planning horizon of AMAN, the ETA can still vary up to 20 minutes. In one SESAR concept for AMAN, such flights are no longer governed by a departure slot but rather by a Target Time of Arrival (TTA), which could be provided by AMAN. It is then up to the AU to try and depart such that it can meet that arrival time [1]. The next chapter will discuss the effects of such a method on accuracy.

By calculating the earliest and latest feasible ETA along their current 3D trajectory, AUs may support the sequence manager in assigning realistic arrival times [13]. The size of the arrival interval may be constrained by operational limitations and by commercial considerations. The difference may be relevant to the sequence manager as the first constraint limits the available options, whereas the commercial constraints only affect potential benefits.

Through SWIM, different ANSPs will also be able to share their requirements on a trajectory—such as an arrival time planned by AMAN—but also their capabilities to provide for such requirements. TBO is planned to allow AUs and ANSPs to develop, monitor, and adjust, a trajectory that realises the planned schedule [1]. While modifications to the trajectory may not always be available (i.e., to avoid conflicts with other traffic), the jurisdiction over a flight—which currently limits the horizon of AMAN—becomes less of an issue as all actors are more aware of the requirements.

Control

Advanced flight guidance systems in aircraft are expected to be able to navigate not only to follow a precise 3D path but also to meet Required Times of Arrival (RTAs) particular at points along that path [1], [39], [40]. Aircraft capable of such 4D navigation can receive the RTA according to the schedule generated by AMAN, and, since the aircraft performs closed-loop control, can be expected to meet that time to a high accuracy. However, the RTA may well need modification to account for disturbances in the schedule (e.g., delay of leader aircraft in the sequence). Furthermore, a combination of aircraft that fly to a RTA and aircraft that are managed by traditional control can lead to separation issues as their speed profiles need not be compatible [41].

Finally, various concepts look at providing additional metering to points further from the airport [13], [42]. While the times over these points at those points are less critical, such metering would improve traffic presentation and the likelihood that a schedule can be realised at the final metering point. Furthermore, required modifications are available through smaller trajectory adjustments, improving efficiency.

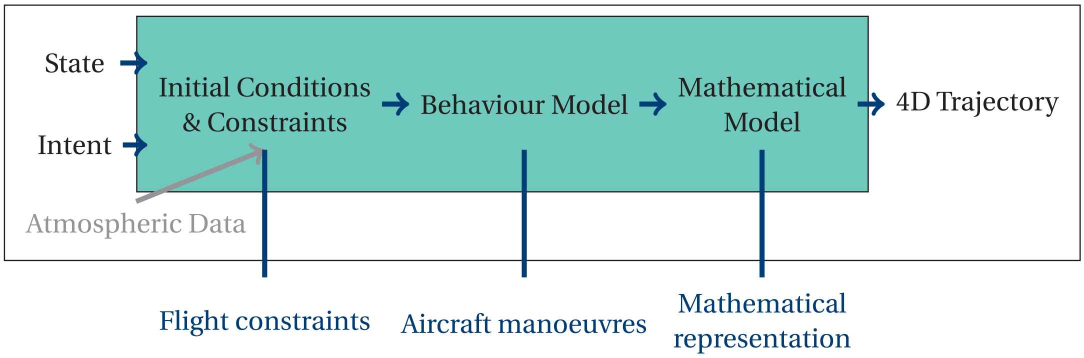


Figure 2.4: Simplified representation of TP concepts. Based on [49]

2.4 Trajectory prediction

Due to the large horizon described in Section 2.2.1, a human operator is unlikely to be able to predict the relative separation between two aircraft at the metering point. This problem is exacerbated by the fact that the two aircraft can be geographically far away from the airport—and from each other—at the moment of prediction (i.e., coming from different directions) [43]. Any decision based on their separation at landing will, therefore, require an automated prediction of the ETA at the metering point using a TP [44].

Various factors may cause errors in an ETA. These include errors in the information used in the prediction, errors made in modelling the flight, and disturbances that deviate the flight from the predicted path [45]–[47]. To deliver the intended benefit, a planning made with such errors requires adjustments as ETAs shift.

The sources and types of errors, combined with their application in the TP determine the accuracy of the TP. To understand the potential effects of errors, this section will explore the components of the TP and how errors may affect the prediction outcome.

Note that the chapter will mostly deal with the *accuracy* of a prediction. Sensor noise and signal resolution influence the precision of a prediction and will be discussed when present, but most are far smaller than the errors.

The components of a TP can be described using the structure proposed by the EUROCONTROL/FAA Action Plan 16 white paper [48]. This section will follow an abstraction of that structure (Figure 2.4) to describe the components and their potential for introducing errors in the prediction. Current AMAN systems have an internal TP or retrieve the prediction from a TP within the ANSP's system. However, the FMS onboard aircraft also includes a TP which could also be used in AMAN as described in Section 2.3.9. Despite this difference in the origin of the prediction, the following description of TPs applies to any application.

2.4.1 State

The state of an aircraft, as relevant to a TP, determines the starting condition of the trajectory:

- Location and altitude: This information is mostly retrieved using radar or approximated using information on the flight's progress. In the first case, the quality of the information derives from the sensors. This accuracy and precision of radar are in order of tens of meters and thus negligible when aircraft cover such distances within tenths of seconds. Flight progress information may include some prediction toward known points and can also approximate progress on the route even though the aircraft is not flying on the route. Both laterally and longitudinally, some error in state estimation may be introduced.

- Time: Both when aircraft are still on the ground and when the prediction is based on flight progress information, the starting time of the prediction can generate a first inaccuracy. Especially when aircraft are still on the ground, variation of departure times of 10-20 minutes is currently common [38], [46]. Concepts such as TTA may help to address this problem (see Section 2.3.9). In most current applications, time is reported in whole minutes. The precision of the information, therefore, does not go lower than this.

- Speed information: Radar information can provide an accurate ground speed. However, airspeeds require either downlinked information from the aircraft or derivation via atmospheric information (primarily wind). The accuracy of atmospheric data influences the latter.

- Mass information: Information on the actual mass is not currently available to any operational AMAN. Mass influences the vertical profile and the landing speed of the aircraft. Errors in the vertical profile may in turn influence the ground speed an therefore arrival time [45], [50], [51].

2.4.2 Intent

Intent is any information on the future path of the aircraft. This information includes constraints, objectives and preferences from both (multiple) ATCOs and the AU. In effect, any information that helps describe the aircraft's planned path, as long as the information is available to the TP. Current AMAN systems have access to a limited set of the actual intent of the flight:

- The filed flight plan will provide a generalised route, cruise altitude, and cruise airspeed. However, such a flight plan is filed at least 3 hours before departure [38]. Due to various reasons, the operator may have opted to change the route, cruise speed, or altitude, therefore introducing an intent error.

- Currently, some ATC systems are aware of instructions provided to inbound aircraft, although few systems can yet include instructions from upstream Flight Information Regions (FIRs).

- In some cases, the landing runway and arrival route of the aircraft are defined by the inbound route (see Section 2.3.5) and therefore available to the ANSP.

Missing intent has to be provided by the system based on assumptions. Therefore, a considerable component of prediction error in modern TPs is caused by incorrect intent information such as speed [45], [50].

SWIM aims to ensure that all existing information is available everywhere. This sharing of information will significantly enhance the accuracy of the intent mentioned in the first two points. The last point is very strongly dependent on the AMAN process itself; while it may be available to the ANSP when predicting the ETA, it may not yet be available to the AU at that stage.

It is foreseen to share such arrival routes through SWIM. However, as runway selection is part of the AMAN process, multiple options may need to be evaluated. The unknown components of intent have to be introduced during the prediction process.

2.4.3 Atmospheric data

Most atmospheric information in AMAN systems is based on forecasts based on earlier measurements. With the forecast comes the accuracy of the forecast model. In less stable atmospheres, differences in temperature but primarily in wind speeds can reduce TP accuracy considerably [52].

Several concepts have demonstrated prediction improvement by direct use of data measured by aircraft that are ahead on the same route. However, these techniques are limited by the number of aircraft on the downstream route [53]–[55]. Similarly, as the AMAN horizon increases, the trajectory prediction will be further into the future, and therefore some form of weather forecasting will still be required.

A final technique is avoiding the use of weather data altogether. As the arrival time is based on the ground speed, calculation using the measured ground speed eliminates the conversion of speeds and inclusion of winds (for example in 4D-Planner [11]). This technique assumes that the wind's effect on ground speed is constant. As the aircraft turns and changes altitude, the effect of wind on the ground speed changes, which may cause considerable speed errors.

2.4.4 Initial condition and constraints

State, intent, and atmosphere information can be defined regardless of the type or purpose of the client TP. However, the information may be of the wrong form or insufficient to define a trajectory in a particular application. Some examples of this problem are:

- Intent information that is not yet available: The route to the runway is not yet decided, or undefined when ATC uses vectors in the TMA.
- The TP model makes specific assumptions that may conflict with the provided intent: certain waypoints are always overflown (a requirement in AMAN when the waypoint is a metering point), even when the flight may bypass those during operations.
- The model may not be able to use information consistently: Temperature at altitude may be available from the atmospheric data, but the system may assume a standard atmosphere. Therefore the vertical temperature profile may need to match the profile assumed in that model.
- Some data can be represented in both the state and the intent, and a choice of the source is necessary. For cruise speed, for example, state information may be ignored instead of intent data or vice-versa.

The output from the constraint model is a description of *what* bounds the future trajectory. Furthermore, this description is compatible with the calculation method for the trajectory.

2.4.5 Behavioural model

Where the constraints model describes *what* trajectory will be flown, the behaviour model describes *how* an aircraft will manoeuvre to achieve that trajectory. Examples of behavioural model parameters are:

- Bank angles used in turns,
- Interception of lateral and vertical paths,
- Delay between instruction and execution of a manoeuvre, and,
- Speed and configuration changes during approach.

Note that this information further defines the future path of the aircraft and could, therefore, be provided as intent. The behavioural model largely introduces intent that is unavailable as a separate input.

2.4.6 Mathematical model

The final component of a TP is the actual trajectory generation. This component implements the intent and behavioural model to a representation of the aircraft. The mathematical model determines how accurate the desired behaviour is represented in the trajectory. In current TP applications, aircraft are typically modeled as either point masses (see [33], [56]–[58]) or points with kinematic properties (e.g., [59]). These point objects are subsequently integrated over time. This type of model needs a detailed description of both the trajectory and the behaviour of the aircraft. Errors or assumptions in intent or behaviour will lead to modelling errors at each time step.

Porretta et al. [60] propose a TP which determines an algebraic solution to the trajectory that satisfies the behaviour and intent. Since AMAN assumes a defined end condition (the metering point), the expected route may well be known. This approach may also include intent defined in time (i.e. RTAs) and may well be more accurate in those cases.

Finally, for sequencing and scheduling, the AMAN only requires an ETA. The intermediate trajectory is superfluous. Functionally, there is no need for accurate intent or behaviour in the intermediate part of the trajectory as long as the resulting ETA is accurate. De Leege [61] makes use of this property by using machine learning to determine a direct relationship between the initial state of the aircraft and the arrival time.

2.4.7 Evaluation: prediction uncertainty

The accuracy in each of the components described above affects the accuracy of the prediction. The uncertainty in the resulting output depends on the uncertainty in the inputs and the prediction horizon. Each type of input and each component will introduce uncertainty to a prediction. Some errors, such as ground speed error, will cause an error on the ETA that grows with the prediction horizon. Others introduce a constant error on the arrival time and are independent of the horizon:

- Errors in the 4D initial state of the aircraft will either introduce a different length path or a different starting time. Most of this error will result in an offset in arrival time which does not change with the length of the prediction. However, when the prediction horizon passes the departure, a sudden, large increase in uncertainty is likely [46].

- Errors in assumed mass may introduce a ground speed error in the climb and descend segments. In these segments, the error is horizon-dependent. Therefore, the effect is constant during the cruise.

- As routes in AMAN end at a known point (e.g., the runway), route intent error can only generate position errors in the intermediate trajectory. Longitudinal errors are affected by both route and speed and are thus horizon dependent [37].

- Atmospheric errors will have a continuous effect on the trajectory, and the error itself will be dependent on time. The impact of atmospheric errors on uncertainty grows with increasing horizon.

- The constraint modelling can both increase errors and bound errors. The previous section provides an example of the first in forcing a route along the metering point when the aircraft will not fly over the point. An example of error bounding is the rule that aircraft are not allowed to deviate more than a certain distance from a route when following that route [62].

- Behaviour modelling makes assumptions on the operator's choices within the constraints. Modelling errors (such as bank angle) are likely to affect the longitudinal profile of the aircraft and thus dependent on the horizon (e.g., each turn adds some error).

- Mathematical modelling (such as point-by-point turn vs constant radius turn [45]) affects the size of the error at each integration step. As such errors accumulate in integration, the error due to mathematical modelling is very likely dependent on the prediction's length. The error will not occur for TPs that instantaneously derive arrival time.

It is commonly recognised that the AU has the most accurate model of their particular flight, including the AU's preferences. If the onboard FMS also has access to the same quality atmospheric data, it is considered likely that a predicted trajectory from the aircraft has the smallest error (i.e., the lowest uncertainty) [37], [39], [63]. However, this is only true if the intent (e.g., runway selection) matches the intent needed in AMAN [37]. Therefore, AMAN trajectories may well be more accurate on the ground as the AMAN may have a more accurate knowledge of the intent of the ANSP for the remaining route.

2.5 Human machine interfaces

Once trajectories are available, HMIs can be used to evaluate the arrival management problem. If the evaluation is manual, the 4D information has to be provided to the human operator in a meaningful way. This section will describe how the information is currently provided and display concepts under development.

Often, the operator will have access to a Plan View Display (PVD) which displays all traffic within the horizon on a map. While this provides the instantaneous geographical situation, representing 4DTs is problematic [64]. Most current AMAN systems, therefore, provide a secondary HMI, which may or may not be included as an element on the PVD.

We will first describe the two most common interfaces available to the planner controller: Schedule lists which aircraft with their arrival times and time lines that show the aircraft on a vertical time axis. The following parts then focus on particular types of information and their presentation in a display.

2.5.1 Schedule lists

These are the simplest form of showing aircraft and their arrival times in a list. These lists are used in the US CTAS - Traffic Management Advisor (CTAS-TMA) and the Dutch, for example Inbound Planner (IBP) [11], [65].

Figures 2.5(a) and 2.5(b) show that the essential information is provided: aircraft identification, ETA, and Planned Time of Arrival (PTA). However, estimating separation requires reading the individual times and mentally calculating the differences. In essence, the display has translated the 4D situation to a digital representation where extracting the relationships between individual aircraft requires mental effort. This need for mental projection has been recognised as a common problem for lists of time-related information [66]. During the years, the time line has, therefore, replaced the schedule list.

2.5.2 Time lines

Instead of providing a digital time, the information is provided in analogue form by showing labels on an axis representing time. This representation enables the understanding of the relation (i.e., separation) between different aircraft at a glance.

All current AMAN systems use a time line [11], [33], [65], [67], [68]. A good example—and one of the earliest implementations—is CTAS-TMA in the US as shown in Figure 2.6 [69]:

- A vertical bar showing absolute time. This bar moves as time progresses, in most cases downward.
- Aircraft attached to their respective arrival time move downward with the time.
- Each aircraft symbol accompanied by the flight identifier.

The vertical distance between two aircraft symbols shows the time-separation. Through this presentation, the interface enables direct comparison of arrival times irrespective of the present location, planned route, or individual arrival time of aircraft. In general, the time line presentation has the following advantages:

- The analogue representation allows for easy visual analysis of spacing between aircraft.
- Short peaks of high demand or “bunches” are recognisable due to the Gestalt principle of *proximity* [70]. Operational experts suggest that such an overview of traffic peaks is often more valuable than the exact schedule.
- The representation provides both the sequence and the schedule. This combination enables concurrent monitoring and manipulation of both.
- Constraints that change over time (e.g., runway availability), can be visualised in the same axis as shown in Figure 2.7.

2.5.3 Time to indicate

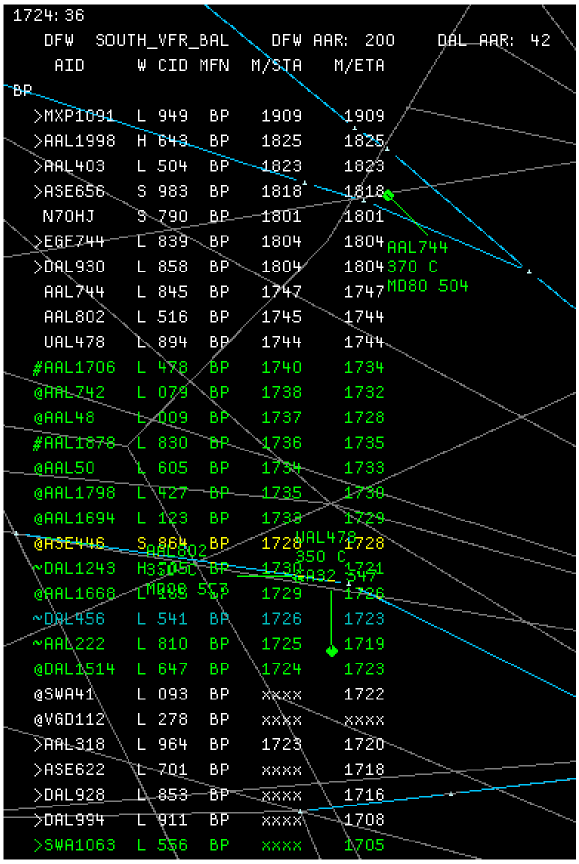
In the AMAN process, three different times are relevant to the sequence manager:

- When the AU prefers to arrive,
- When the aircraft is expected to arrive based on its known intent: the ETA,
- When the aircraft is planned to arrive after optimisation by the AMAN process: the PTA.

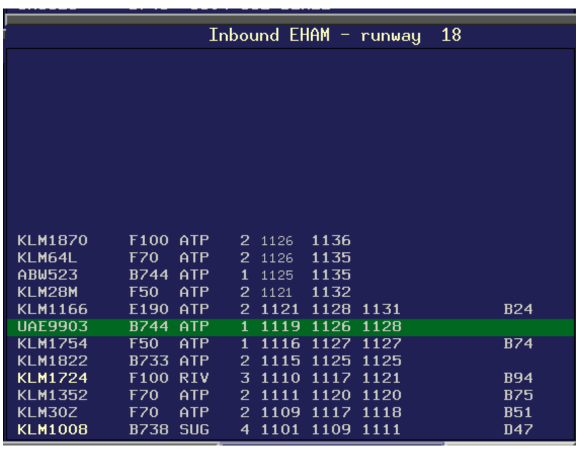
If not constrained by safety, capacity, or workload, the optimal arrival time is the time at which the AU prefers to arrive (see Section 2.2.1). As long as the aircraft’s trajectory is not yet constrained by ATC instructions, a trajectory may be defined that complies to fixed constraints such as routing and is optimal from the AU’s perspective. This time may well deviate from the original STA as defined in the flight plan. For example, some AUs may choose to increase speed to recover a delay, while others may choose to maintain the most fuel-efficient profile and accept the delay.

An aircraft label can only be attached to one location (i.e., time) on the time line. Therefore only one of the three times (STA, ETA, or PTA) can be indicated using the visual representation. This limitation introduces a challenge in selecting the time to which the label should be connected to and how the other two times can then be presented.

When a PTA is communicated to the upstream controllers or the AU, they will adjust the trajectory to meet the PTA. It could be argued that the difference between ETA and PTA becomes transparent to the sequence manager: In his or her view, the trajectory between the current state and the metering point may be undefined, but the controller expects aircraft at the metering point on the PTA. In effect, the PTA now becomes ETA.



(a) CTAS-TMAas overlay on the radar screen [65]



(b) IBP [11]

Figure 2.5: Examples of schedule lists.

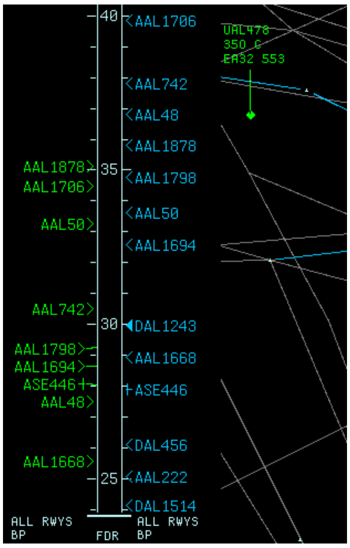


Figure 2.6: Time line representation as used in CTAS-TMA [65].

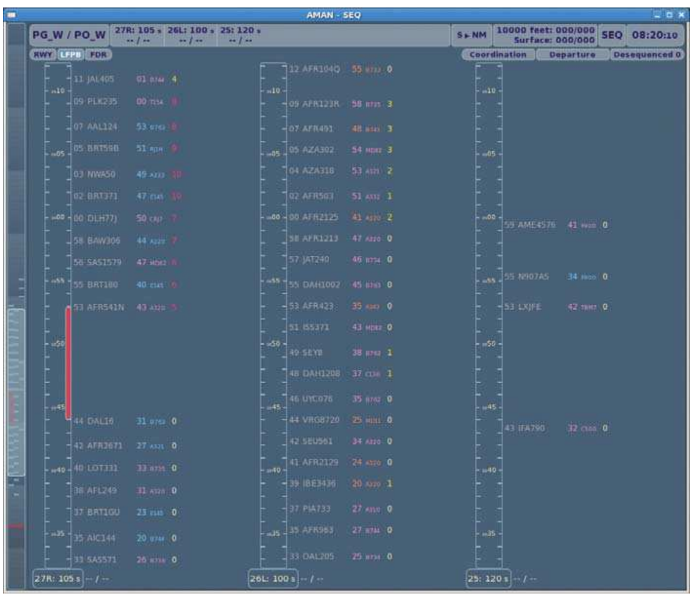


Figure 2.7: MEASTRO time line showing a runway closure as a red bar over a time interval [67].

As the trajectory may not be adjusted immediately (e.g., the controller plans to provide the required speed reduction for a delay at a later time during the flight), the Time to Loose/Gain (TTL/G) may only gradually reduce to zero. This gradual decrease also means that the actual ETA of the trajectory may still be relevant to the sequence manager as it influences the optimal arrival time for possible replanning. The first question, therefore, has no definite answer. However, all current operational systems place the label at the PTA.

To provide the other times, CTAS-TMA uses two sides of the time line. The left side shows the aircraft at their ETA, and the right line shows the labels according to the plan. In other representations, the TTL/G is shown as a number or an indicator in the label as in DFS's 4D Planner [11]. None of the analysed systems indicate the STA or the AU's preference at the moment.

2.5.4 Presenting separation

To determine whether two aircraft are sufficiently separated, the sequence manager has to know the expected separation and the required separation of each aircraft pair. At the metering point, the time line interface directly shows the expected or planned separation in time—depending on whether the display uses ETA or PTA, respectively. For distance-based spacing, separation has to be derived from the time interval and the trailing aircraft's speed over the metering point. Therefore, determining the planned separation in time requires knowledge of the aircraft's properties and the headwind at the metering point. Determining the separation distance then required mental effort to calculate.

While technically feasible, no operational system shows the required separation between aircraft on the time line. The only concepts with a a display of the separation requirement are part of research simulations [34], [71]–[75]. The literature does not provide any evidence against such an indication of the requirement. In discussions on the subject, operational ATCOs have argued against displaying the separation requirement as it may lead to spending too much effort on exactly achieving the desired separation where some variation is currently considered acceptable. This drive for accuracy may subsequently lead to an unnecessary increase in task demand. This proposition has not been validated in experiments.

2.5.5 Geospatial context

Section 2.3.7 explained that not all schedules might be achievable, as the resulting trajectories may conflict in the intermediate segment between the current position of the aircraft and the metering point. Since both schedule lists and time lines are based on the time at the metering point only, these displays cannot present the spatial relationship. In these cases, automation either detects and resolves this conflict without involving the sequence manager (such as CTAS - Efficient Descent Advisor (CTAS-EDA) [76]) or requires the controller to judge the feasibility of the schedule based on the situation provided on the PVD.

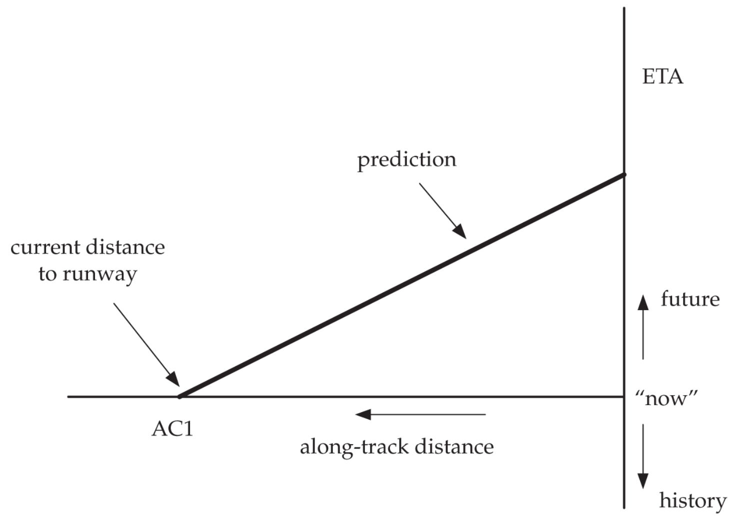


Figure 2.8: TSD concept.

At the Delft University of Technology, various studies have looked at incorporating spatial relations with a time line using a time-space representation. In these HMIs, the predicted along-track spacing in time is shown as a function of distance to the metering point (See Figure 2.8) [72]–[75]. This representation enables the visualisation of separation and separation requirements both at the metering point and on the full trajectory toward that point (Figure 2.9). Furthermore, such an indication of separation requirements can be used to analyse the effect of different sequences. However, preliminary experiments do suggest that human operators have difficulty recognising the potential benefit of a sequence switch from the display [77].

The time-space representation visualises separation over the common path of aircraft. In doing so, it disregards the geographical and vertical position of the aircraft. This simplification also reduces the ability to represent traffic that does not share that common path. Visually connecting information on the TSD and PVD can partially address the first problem [75]. The latter may partially be addressed by representing the conflicts due to other traffic onto the TSD [74], [75].

A notable argument against an HMI such as the TSD is the need for a new display. This extra screen requires physical space on the console and a division of attention of the controller between separated screens [34], [75].

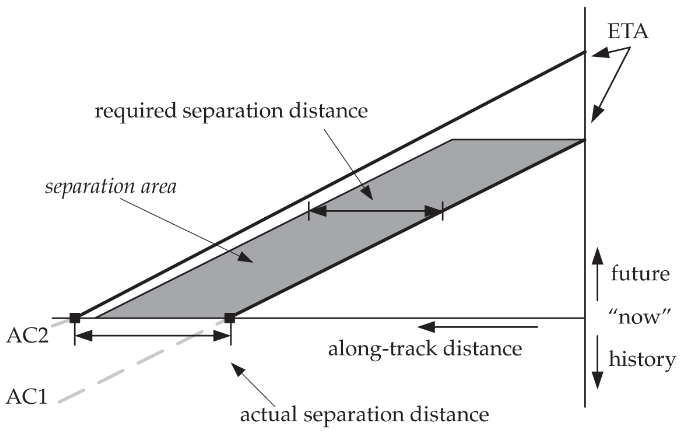


Figure 2.9: Representation of separation and separation requirements both at the time line (right) and during the path towards arrival.

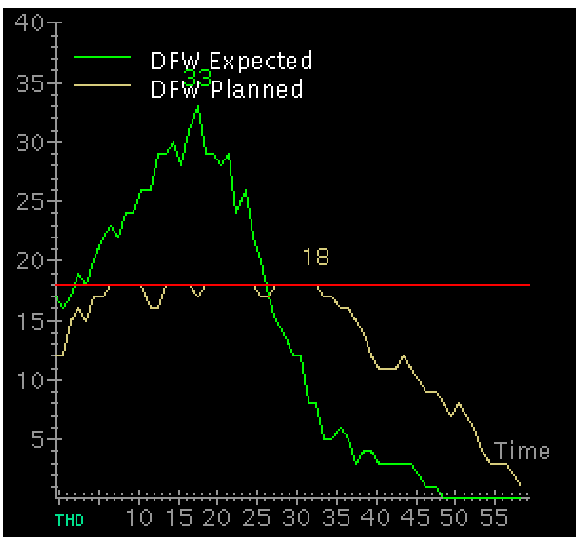


Figure 2.10: Traffic manager's load graph in CTAS-TMA. The green line shows the required capacity, the red line shows the available capacity and the brown line shows the throughput after planning. Source: NASA [65]

2.5.6 Presenting capacity

When the airport cannot provide the capacity to meet the demand, the capacity needs to be increased (e.g., adding a runway) or demand needs to be reduced by delaying aircraft at the back of the sequence. To make such decisions, the sequence manager must determine whether the current sequence of inbound aircraft exceeds capacity.

When the DSTs shows the required separation, it may provide means to try and fit the offered sequence to the metering point. However, the consequence of a shortage of capacity is only presented as large a delay on the last aircraft to arrive once that last aircraft is scheduled (i.e., after the process has been applied). Very few systems provide early indications of capacity shortage.

In CTAS-TMA, a graph supports an early understanding of a capacity problem by showing the amount of inbound aircraft per time interval and the available capacity [65]. The graph (See Figure 2.10) furthermore indicates the effect of the current plan on capacity. It provides an overview of capacity and demand, expressed in the number of aircraft only. Section 2.2 describes the number of aircraft as one of the factors that determine the controller task demand, which is one of the limiting factors on capacity.

However, the number of aircraft is not the only factor, as route structure, weather, and team competencies may also affect the workload. The PHARE project has demonstrated a display that provides information on throughput as well as expected workload [78].

2.6 Automating the process

When the sequence manager is presented with ETAs, a large part of their task would be to manually space the individual arrival times according to the separation needs. Especially when inbound traffic is well regulated beforehand (e.g., by the NM), it is likely that there is little overlap between arrival times. This fact is also recognised in the development of automated schedulers: A so-called First-Come-First-Served (FCFS) approach often is close to the optimal solution [8]. The management task would become primarily administrative in shifting aircraft small amounts of time to remove minor errors. At the other extreme, real optimisation to all objectives is often considered too complicated for a human operator [10].

This section describes operational and proposed concepts for sequencing and scheduling automation. In doing so, it will try to follow the process described in Section 2.2: Runway planning, sequencing, and finally, scheduling. While generally performed as part of an integrated process, the following sections separate sequencing from scheduling as the first requires considerably more advanced functionality.

2.6.1 Runway Planning

Runway planning is a collaboration between the airport operator, departure planning, and arrival planning. Some systems provide algorithms for optimising departures and arrivals for a given runway plan (e.g., Osyris [79]). Future integration of AMAN and Departure Manager (DMAN) is likely to improve this process [1].

The decision on a runway plan is currently based on procedures, availability, weather, and demand in departures and arrivals. The Netherlands Aerospace Centre has developed a decision support system that informs the user on the available runways, including the likelihood of weather conditions on the availability [80]. Based on the tools, the system predicts the most likely configuration. However, the latter should not be seen as a decision support system as it predicts the controller's decision based on historical data. Using the system as decision support in selecting runways would create feedback [81].

2.6.2 Scheduling

Basic runway allocation rules, such as the examples in Section 2.3.5, can be implemented in an automated system. Several operational systems currently implement this strategy [68], [82]. However, that section also explains that some strategies may not be as trivial. For example, when controllers create "trains" of aircraft from a single direction.

Once aircraft are assigned to runways, PTAs can be determined. The spacing between two aircraft can be defined as an adaptable parameter, either constant, depending on the wake turbulence categories of the pair, or a combination of both. Therefore, the next aircraft's arrival time is at the minimum of the next available slot or its ETA, whichever is later. This approach is often named FCFS. Most systems automatically assign aircraft to these slots using such basic rules [11], [69], [82].

In scheduling automation, the difference between separation in time and separation in distance is essential. When time is the spacing criterion, the following aircraft can land at the ETA of the previous aircraft plus the required interval. When distance is used, the aircraft's ground speed has an effect: Once the leading aircraft has passed the measurement point, the trailing aircraft has to cover that distance at its own ground speed. Especially in landing—where speeds depend on aircraft type and mass—the interval in time may vary considerably. The accuracy of the interval thus relies on the knowledge of the landing speed and the local wind conditions (see Table 2.2).

Assigning delay does not necessarily require any trajectory prediction beyond determining the ETA. While the delay can grow to unrealistic values, no additional aircraft can land if there is insufficient capacity; The capacity objective takes precedence over efficiency. In contrast, very few concepts evaluate the possibility of time advance/frontloading (e.g., CTAS [69]). Time advance requires the calculation of the feasibility of the new arrival time (see Section 2.3.6), which may require further trajectory prediction.

2.6.3 Sequencing

Section 2.2.2 explains the effect of wake-turbulence categories on required spacing. Some systems automatically optimise the sequence based on these categories [11], [58], [82]. Sequencing automation relies on trajectory prediction to evaluate options on which aircraft to land first and which to land second such as:

- Physical possibility of the switch (i.e., can the trailing aircraft fly fast enough to overtake the leading aircraft),
- Benefit of switch against the efficiency of the two profiles,
- Maximum acceptable delay assigned to a single aircraft (equity between operators),
- Safety considerations of a switch (e.g., two aircraft already following each other may violate separation criteria during overtaking).

2.7 Evaluation of human-machine collaboration in AMAN

Section 2.2 described the AMAN process, explaining the objectives of arrival management and how these can be achieved. Sections 2.4 until 2.6 described the required underlying TP and the uncertainty of its outputs, the HMIs available for human operator to support the process, and the potential for automating the process. This section will discuss how these components affect the ability of the human operator to ensure an optimal arrival process. The first part will address the potential effects of prediction uncertainty, the second part analyses how the different HMIs support the sequence manager's decision process, and the last part looks at the potential collaboration between the human operator and the supporting automation.

2.7.1 Effects of prediction uncertainty

Section 2.4 explained that both uncertainties in the input and modelling choices will limit the accuracy of the ETAs used in making an arrival schedule. Erroneous ETAs could lead to modifications in the arrival plan at a later stage in the approach, which may reduce the performance toward the objectives. These later modifications also increase the workload of both the sequence manager, who has to modify the plan and the upstream executive controllers, who have to execute the plan.

The effects of uncertainty

One of the reasons that current operational AMANs have a limited horizon (see Section 2.3.3) is that uncertainty typically becomes too large for the prediction to be of any practical benefit as the resulting plan is too likely to require modifications [32], [83]. Even if new systems such as SWIM provided state information on a flight at that horizon, it would not enable developing a useable plan (e.g., MAESTRO considers a sequence to be stable at 15 minutes before the metering point [11]).

However, the different factors that influence uncertainty are not constant. Section 2.4 gives the example of aircraft still on the ground versus those in the air. Another example would be stable weather versus unstable weather—where wind errors are likely to be larger.

The actual uncertainty could well vary from flight to flight and from day to day. Underestimation of the uncertainty could lead to a plan based on too high uncertainty, which is therefore likely to need later modifications, reducing its benefit. Overestimation of uncertainty could delay the planning unnecessarily and thus miss an opportunity to get additional benefits from using AMAN.

Using uncertainty

Making predictions more accurate reduces uncertainty and its effects. Concepts such as SWIM and TTA could help reducing uncertainty on intent. Sharing and executing 4DTs and using RTAs could make the uncertainty in ETA dependent solely on the aircraft's capability to meet it. However, neither option is free from uncertainty as, respectively, it still requires predictions, or the uncertainty is defined an acceptable—but non-zero—error. Especially in high density airspaces, the need for separation may limit ATC to provide for a 4DT that meets the RTA [2]. Finally, in the transition to these operations, aircraft capable of 4D navigation will share the airspace with aircraft that are not, resulting in a different uncertainty per flight.

The human operator will have to judge the quality of the prediction by understanding the automation (e.g., selected routes), the resulting ETA, and possibly knowledge of the nature of the inputs (e.g., unstable weather). Some inputs may, however, be transparent to the operator. They can not evaluate the effect on accuracy (e.g., assumed aircraft speed). Research and operational systems have demonstrated the possibility of automatically estimating the uncertainty of predicted trajectories [46], [83], [84]. This information could help decide whether a planning decision is warranted or whether further modifications are too likely.

2.7.2 Available information

The objectives described in Section 2.3.1 are the high-level goals of the AMAN operation. The HMI should ultimately provide support to achieve these objectives.

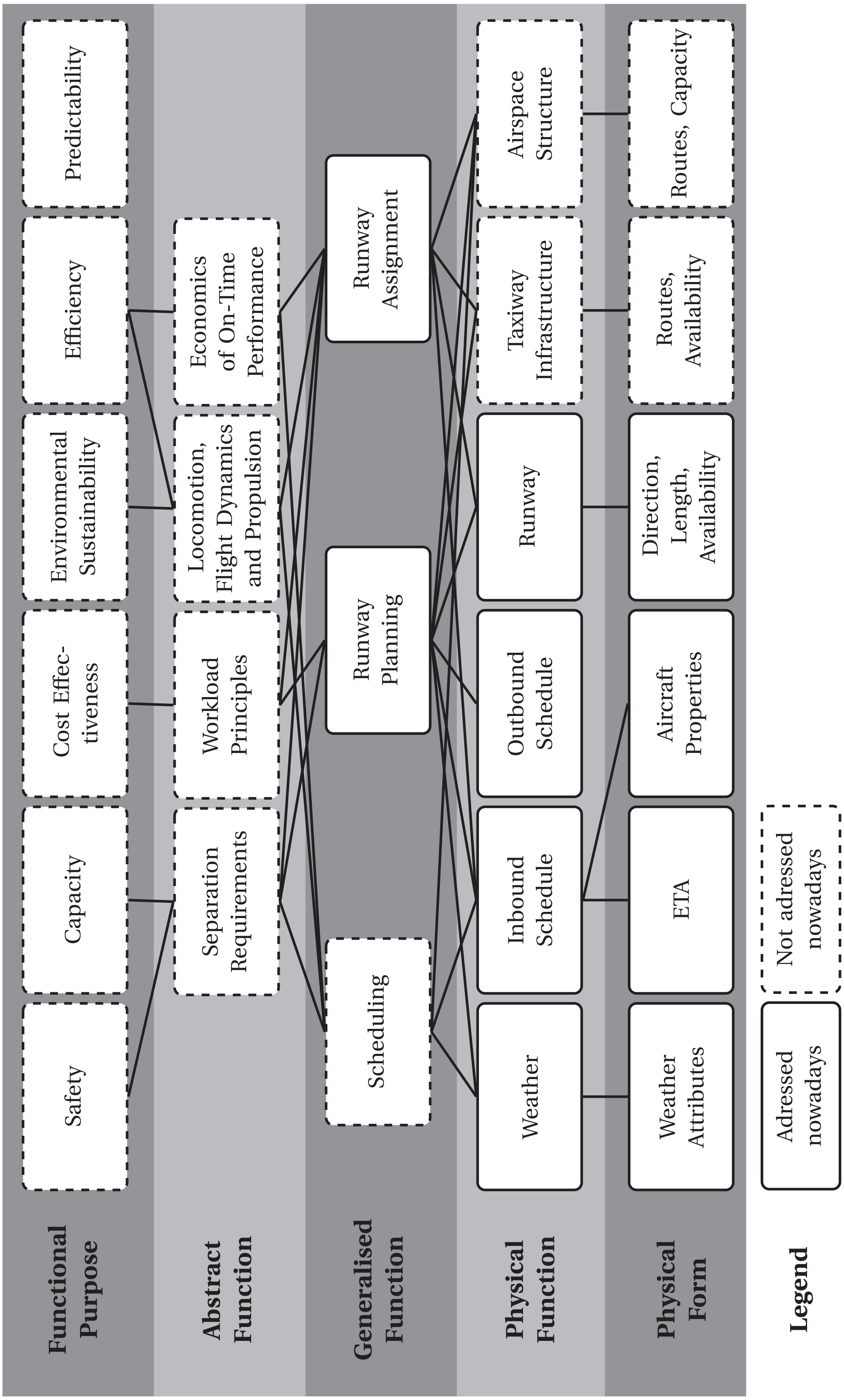


Figure 2.11: Abstraction hierarchy of the AMAN planning problem.

In their concept of Ecological Interface Design (EID), Vicente and Rasmussen [85] propose that effective support for complex systems can be provided by supporting the expert user in understanding the relations between high-level objectives and the detailed attributes of the system. Their *abstraction hierarchy* provides a means to describe why certain details need to be changed to achieve the objectives, or how certain objectives can be achieved by modifying details [86].

This section will not try to develop the complete abstraction hierarchy of the arrival management problem. Instead, it will work backwards from the items available in the different systems, placing them in the means-end relationship of the schema. Figure 2.11 shows a high level abstraction hierarchy based on the objectives from Section 2.3.1.

Safety

The primary safety objective for AMAN is to facilitate separation and spacing (in-trail separation). While both describe a distance between aircraft, the distinction helps in recognising evaluating actions that prevent risk due to wake vortex encounters and those actions that prevent collisions between any aircraft.

Both schedule lists and time lines indicate the PTA, which is the desired end-state of the arriving aircraft. The in-trail spacing itself is represented at the abstract function level. Section 2.5.3 notes that the ETA is only directly available in a few systems; expected spacing is therefore not directly presented. Finally, none of the operational HMI concepts indicates the required spacing.

To determine the required spacing in time, the human operator will have to spend mental effort to assimilate data on the aircraft type and its expected landing speed. More advanced automated systems calculate the spacing required between each aircraft pair which accounts for the varying effects of wake vortices [87]. Therefore, the automated sequence can be said to provide the controller with information on the required spacing. However, the resulting information hides the underlying principles. Analysing the benefit of a switch in sequence would, therefore, still require mental effort.

Finally, no automated planners or operational planning interfaces calculate or provide information on geographical separation from other aircraft or weather obstacles. This problem only is partially addressed in the TSD concept (see Section 2.5.5) and in some automated solution advisory systems such as CTAS-EDA [76]. For the planning function within AMAN, the separation component of safety is therefore only supported as spacing. Simulations do indeed suggest that such concepts may improve delivery accuracy but provide limited improvement on separation risk [32].

Capacity

AMAN is needed when a capacity shortage exists, making capacity the primary purpose. At the functional purpose level, only the graph in CTAS-TMA (Figure 2.10) provides information on demand, available capacity, and realized capacity. Section 2.5.6 explains that the required delay for the last aircraft can indirectly indicate the capacity shortage. However, that does assume that the sequence and separation have already been optimised for maximum use of the available capacity.

At the level of separation, capacity and safety share the same relations; The description for the safety objective also holds for the capacity objective. For capacity, automated systems may make additional switches in the sequence based on WTC to improve total throughput. However, the potential effect of switches is not available in any operational system. An indication of required separation could potentially provide the information, although current HMI concepts have not yet been able to demonstrate that capability (Section 2.5.5).

The two paragraphs above consider the planning of aircraft to the available capacity. Assigning aircraft to the available capacity should be performed in combination with the planning of that available capacity (e.g., number of runways to use). Figure 2.7 showed how runway availability can be visualised on a time line. This technique can be used for temporary closures of runways but also for planning a mix of arrivals and departures. However, the need for a common point in the different representations does limit the ability to represent multiple runways on a single line (see Section 2.5.5). It is, therefore, hard to plan for the use of the capacity without first assigning aircraft to a runway. Yet, the assignment of aircraft to a runway may affect the capacity due to the wake turbulence categories of the aircraft.

Efficiency

Since efficiency derives from the cost of fuel and the costs of not arriving on time, the most efficient trajectory is solely determined by the needs of the AU. If the AU can determine such an optimal trajectory, and there are no other constraints on the route, the ANSP could provide that most efficient trajectory (see Section 2.5.3). This is the concept of the Reference Business Trajectory (RBT) as envisaged by SESAR. Note that in current operations, AUs already make such choices; for example, by adjusting their cruise speed based on the expected delay.

Few controllers will issue instructions to change the trajectory of an aircraft from its preference if there is no need. However, other than the current speed of the aircraft, the current systems do not provide information on the operator's preference. Such support will require more data from the AU regarding the preferred trajectory of the flight when approaching the destination (see Section 2.2.1).

Cost effectiveness

As explained in Section 2.3.1, the cost effectiveness of air traffic services strongly depends on the amount of staff needed. The amount of staff needed subsequently depends on the total amount of workload and the maximum workload for a single operator. Workload limits are currently imposed as capacity limits by the AMAN operator. Automated support will require modelling of the controller's workload (see Section 2.2.2). Similarly, an interface that supports workload management can provide such properties at the lowest levels (e.g., the location and routes of aircraft) but will require modelling for higher levels (i.e., to indicate the expected workload based on the plan). Therefore, the current HMIs do not support strategies that reduce workload by other means than reducing the flow rate.

Environmental sustainability

None of the evaluated concepts include explicit components relating to either emissions or noise. As mentioned in Section 2.3.1, AMAN may support CDOs by reducing the need for route deviations close to the runway. Such fixed routes improve the ability to perform CDOs reducing both noise and emissions in the approach [26], [35].
Secondly, it could be argued that the objective to minimise fuel consumption for the AU's benefit shares much of its factors to make the flight environmentally efficient. However, on-time performance may make the flight efficient from the AU's—economic—perspective but increase the fuel used to achieve those times. Therefore, efficiency and environmental sustainability should be supported separately.

Predictability

AMAN itself should improve predictability by providing tools to timely and effectively manage capacity and demand. A good collaboration between arrival planner, automation, and executive controllers will support the predictability objective. Therefore, predictability can be seen as a consequence of design, rather than a design feature.

Another way of addressing predictability is actively using or controlling robustness in the arrival schedule. The first may be achieved by using knowledge of the uncertainty, as discussed in Section 2.7.1. The second involves reducing uncertainty by metering through progressively smaller windows as described in Section 2.3.9.

2.7.3 Automated processes

SESAR specifies a clear role for human operators to be in command as overall system managers despite high levels of automation. In doing so, SESAR acknowledges that the human operator can only be responsible for automation if the functioning of the automation is sufficiently clear [1]. It is worthwhile to evaluate the current HMI concepts, in which a human sequence manager is responsible for the resulting arrival schedule.

This analysis is performed using Billing's guidelines for human-centred automation in aviation [88, pp. 237–260]. While not a definitive list of criteria for developing all automation, the list provides several helpful points to focus on.

Note that Billings has assembled the list for aviation in general. Several guidelines are more applicable to cockpit automation. This section focuses on the subset most applicable to ATC and DSTs.

Restricting options

According to Billings, automation should be designed to support all strategies that the human may use to allow the human to be in command. Furthermore, automation should make those strategies easier to manage.

The current HMIs are all based on assigning PTAs to aircraft. However, the absolute arrival time is not always relevant to the objectives. Section 2.3.5 for example describes a strategy of making "trains". Deliver all aircraft from one sector at a regular interval to one runway while all other aircraft are directed to another. This separation reduces the number of merging aircraft, which may reduce the monitoring task load of the downstream controller.

In this example, a deviation from the absolute arrival time may be of little consequence as long as the relative separation is sufficient. However, the AMAN system can only communicate absolute arrival times to the upstream controllers. As the automation does not support this technique, the sequence manager may abandon the DST and bypass the automation by direct communication with upstream controllers.

Such bypassing of the automation has another effect: the automation is no longer aware of the sequence manager's plan. As a result, the system will not be able to support the strategy, even if the controller plans to return the absolute arrival times at a later point in the planning. To keep the automated system up to date, the human operator now has to perform the administrative task of entering appropriate estimates of absolute arrival times for the aircraft for which these times are irrelevant.

Another example of potentially restricting options is the need for trajectory prediction if advancing an aircraft to an earlier time is considered (see Section 2.6.2). While a reasonable delay may always be considered feasible, speed limits and route structure may limit how much earlier aircraft can arrive. Most current automated systems will, therefore not consider advancing in the optimisation of a schedule.

Involved operator

Billings argues that a human controller needs to be actively involved or “drawn in”. This does not mean that tasks need to be performed by the human, but the operator should perform meaningful tasks when managing the automation. Most AMAN systems automatically generate an initial schedule based on simple rules (FCFS, route-runway combinations, standard separations). Some systems go further and optimise the schedule for wake turbulence separation or specific runway preferences (see Section 2.6.2). The ATCO is not actively involved in that procedure but may modify the schedule after the assignment of the initial PTAs.

Especially when using straightforward rules, it could be argued that assigning an initial PTA does not prevent the operator from being involved in the planning process. This step merely builds a base to work from and even prevents the controller from performing routine, less meaningful tasks. On the other hand, such simple rules do not account for other factors in planning an arrival time. It may well be that, when these additional factors are taken into account, another planning provides better performance toward the complete set of objectives. By providing an initial planning, the automation may mask other (more suitable) options from the controller.

Awareness of performance

To be able to monitor an automated system, and to take control when the automation fails to perform its tasks, the human operator needs to be aware of what the automated system is doing. The previous paragraphs argue that simple automated scheduling could be applied effectively to reduce routine tasks. The operational automated schedulers place the inbound aircraft at a suitable PTA. However, if the original ETA is not available anymore, it may be uncertain whether automated scheduling has been applied (i.e., whether the ETA differs from the PTA).

Scheduling automation is affected by the objectives that define the optimum sequence. As discussed earlier, an initial scheduling based on simple rules may be observable for the operator. However, once more complex criteria are introduced, the current systems provide little indication of the factors that underpin a particular decision.

Predictable behaviour

To be able to monitor and trust the automation, the human operator has to understand how the automation will respond to the situation and how the response will develop over time. Section 2.4 explains that a human cannot accurately predict the trajectory over a typical AMAN horizon. Humans will, therefore, be poorly able to determine the correctness of the ETA unless they are also informed of the reasons why a certain action was taken.

A common complaint among AMAN users is the lack of understanding of the automated solution[11], [34], [89]. At this point, the difference in representation and objectives influence predictability. There may be constraints that are not detected—or taken into account—by the automation. A good example is the consideration of workload by the operator while the automated system optimises for capacity only. If the workload is not a factor, an automated system may well generate a feasible, optimal sequence that is far too complex to be executed by the—human—controllers.

Relying on reliable systems

Billings states that humans will rely on reliable automation. Section 2.7.1 described how underestimating uncertainty may lead to late—and possibly inefficient—corrections of the trajectory. If a human operator is used to an automated process planning reliably at a certain horizon, an unexpected error could reduce planning performance. Information on uncertainty, as suggested in Section 2.7.1, could prevent such effects of false trust.

The case for automation

This evaluation of automation in AMAN systems lists several problems. However, the conclusions do not preclude using automated scheduling or sequencing. This section even argues why basic automated scheduling might prevent the human operator from having to perform a routine, administrative task of keeping all PTAs matching to the chosen strategy.

What is needed is effective communication between the human operator and the automation [90]. This communication should provide:

- Information on the application of an automated decision,
- Information on the reasons for a certain automated decision,
- Information on the reliability of an automated decision,
- A means to keep the automation informed of the operator's intent, in particular of the operator's strategy.

2.8 Conclusions

To balance the flow of inbound aircraft and the capacity at airports, more and more ANSPs use AMAN systems. These provide support to air traffic controllers in deciding on runway configurations, scheduling, and sequencing inbound flights, to optimise capacity, flight efficiency, and predictability, while preventing excessive workload for the downstream controllers.

In the coming years, ATM is expected to move from a sector-based tactical form of control to a more strategic approach based on developing capabilities to accurately predict, share, and execute 4DTs. In this form of control, aircraft will adhere to a strict lateral and vertical route wherever possible, and time will be added as an explicit control variable. AMAN will play an increasingly important role in planning the arrival times into an airport.

This chapter analysed the human-machine collaboration for the air traffic controller responsible for creating and monitoring the arrival plan (the sequence manager). Three areas were highlighted: the effects of prediction uncertainty on the horizon at which AMAN can be used effectively, the information available in the sequence manager's HMI, and the potential effects of including automated sequencing and scheduling algorithms. This section will discuss the main conclusions of these three elements.

2.8.1 Uncertainty

One of the key requirements on AMAN within SESAR is to extend its horizon. The planning horizon in current arrival management processes is mainly determined by the availability of information on the inbound flight, the ability to influence traffic, and the stability of the predicted arrival times. SESAR foresees the first two problems to be overcome by sharing trajectory information, in which the destination ANSP will influence the target arrival time and all upstream ANSPs cooperating in facilitating that time.

Currently, the AMAN process starts between 20 and 40 minutes before landing. Human sequence managers cannot estimate arrival times at these horizons with sufficient accuracy. Current systems, therefore, use automated TPs to predict ETAs from which to start the planning process. Even with such automation, however, the accuracy of the prediction is limited and decreases when extending the prediction horizon as disturbances have a higher likelihood of occurring and get more time to influence the trajectory. To the human sequence manager, an effective planning horizon will exist beyond which the human operator no longer trusts the automation to provide sufficient support. The lack of reliability of a solution based on an inaccurate prediction is one of the leading reasons for limiting the AMAN working horizon.

With the advent of better trajectory prediction and future 4D navigation capabilities, new technologies can improve the benefits achieved through AMAN systems. While an increase in prediction capabilities will result in more accurate trajectories, there will always be a degree of uncertainty in the solution. At a certain horizon, new types of uncertainties may come into play. An example is the accuracy of the estimated take-off time of an aircraft still on the ground, which may prove too large to allow effective arrival planning.

The human perception of uncertainty will be based on experience and not on the actual uncertainty at a given time, as not all the factors influencing uncertainty may be known. The actual uncertainty is far from constant, as it may vary due to, for example, weather, flight phase, or upstream traffic density. Even more so, during the transitional phase—which due to the long life cycle of aircraft, may take 15-20 years—aircraft with 4D navigation capability and, therefore, high accuracy will share the airspace with aircraft with much lower accuracy for a considerable amount of time.

2.8.2 Information

Most AMAN HMIs are based on a time line showing an aircraft label connected to its planned arrival time on a vertical axis. This interface provides a good overview of separation in time. However, as the label can only attach to one place on the line, other relevant times (such as the AU's preferred arrival time and the ETA) have to be presented in another way.

The information on the interface provides limited support toward managing capacity and safety objectives only. The absence of direct indications of the required separation makes it more difficult to optimise the separation manually. Automated initial spacing may provide cues in the required spacing but hides the underlying principles that govern it. In general, information on the possibilities of achieving higher level performance toward the objectives—for example, total capacity and demand—is seldom directly available. Such information could improve the operator's performance in managing arriving traffic.

AMAN systems are intended to optimise the inbound flow to several objectives. However, the current interface design only supports separation management. Other objectives, such as controller workload, are managed through AMAN but only by applying the operator's knowledge and experience. Especially when other objectives are to play a role (e.g., the AU's preferences), adequate support will require suitable information regarding all objectives.

The focus of the currently operational time lines is on the schedule at the metering point. While aircraft may be separated in time at that point, the systems provide no information on the separation in the intermediate trajectory. Planning solutions at the metering point may lead to conflicts in the upstream traffic. At the same time, such a single metering point limits the use of a single time line in a scenario with multiple runways. Displays cannot easily represent flights toward different points in the same dimensions.

2.8.3 Automated sequencing and scheduling

Most current sequencing and scheduling algorithms apply basic rules to assign traffic to runways and to set the required separation time. In doing so, the automation reduces the sequence manager's task load by eliminating a largely administrative function.

However, the strategy applied in such algorithms does not always match the strategy of the sequence manager (such as assigning traffic to particular runways to limit downstream controller workload). In these cases, the automated algorithm cannot provide adequate support. It might, at times, actually require extra work from the sequence manager to ensure that the system has the correct information to provide further support.

2.8.4 Recommendations

Recent developments, in both experiments and operational ATC, have demonstrated that the uncertainty for a trajectory may be calculated and used in operational decision-making [78], [91], [92]. It might be worthwhile to investigate whether information on the uncertainty can improve the performance of AMAN. In such a concept, the operator or the automation will use the actual uncertainty as part of the process. First, this concept will require an algorithm to predict the uncertainty of an arrival time. Secondly, methods have to be developed to communicate that uncertainty to the human operator or include the uncertainty in automated solutions.

To enable support toward all objectives of AMAN, improvement in HMIs should be investigated that extends beyond presenting the ETA or PTA. The different factors that determine an optimal schedule need addressing to do so. Frameworks such as the abstraction hierarchy may help determine which information could support sequence managers in including a broader set of constraints when making a decision.

If the human operator is to manage automated arrival management, the communication between the operator and the automation has to be improved. Key aspects are information on the automated process, information on the reasons for automated decisions, the ability for the operator to communicate their intent, and finally information on the reliability of the automated decision. The latter could be pursued in combination with information on prediction uncertainty.

2.9 Bibliography

[1] "SESAR concept of operations at a glance ed 02.00.00," SESAR JU, Tech. Rep., 2011.

[2] "D3 - The ATM target concept," SESAR Consortium, Tech. Rep., 2007.

[3] JPDO, *NextGen integrated work plan: A functional outline*, 2008. [Online]. Available: `http://www.jpdo.gov` (visited on 03/01/2013).

[4] T. J. Callantine, E. A. Palmer, and M. Kupfer, "Human-in-the-loop simulation of trajectory-based terminal-area operations," in *International Congress of Aeronautical Sciences (ICAS)*, Nice, France, 2010.

[5] P. Jorna, D. Pavet, M. van Blanken, and I. Pichancourt, "PHARE ground human machine interface (GHMI) project," EUROCONTROL, Brussels, Belgium, Tech. Rep., 1999.

[6] R. Göritzlehner, C. Borst, J. Ellerbroek, C. Westin, M. M. Van Paassen, and M. Mulder, "Effects of transparency on the acceptance of automated resolution advisories," in *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, San Diego, CA, USA, 2014.

[7] J. E. Beasley, M. Krishnamoorthy, Y. M. Sharaiha, and D. Abramson, "Scheduling aircraft landings? The static case," *Transportation Science*, vol. 34, no. 2, pp. 180–197, 2000.

[8] X.-B. Hu and W.-H. Chen, "Receding horizon control for aircraft arrival sequencing and scheduling," *IEEE Transactions on Intelligent Transportation Systems*, vol. 6, no. 2, pp. 189–197, Jun. 2005.

[9] L. Bianco, P. Dell'Olmo, and S. Giordani, "Scheduling models for air traffic control in terminal areas," *Journal of Scheduling*, vol. 9, no. 3, pp. 223–253, Jun. 2006.

[10] T.-C. Wang and Y.-J. Li, "Optimal scheduling and speed adjustment in en route sector for arriving airplanes," *Journal of Aircraft*, vol. 48, no. 2, pp. 673–682, Mar. 2011.

[11] N. Hasevoets and P. Conroy, "AMAN status review 2010," EUROCONTROL, Brussels, Belgium, Tech. Rep., 2010.

[12] ——, "Arrival Manager - implementation guidelines and lessons learned," EUROCONTROL, Brussels, Belgium, Tech. Rep., 2010.

[13] A. Barff, B. Favennec, P. Conroy, *et al.*, "SESAR P05.06.04 - D28 - Preliminary OSED ed. 00.01.01," SESAR Consortium, Tech. Rep., 2012.

[14] Boeing, *Airport noise and emissions regulations*, 2012. [Online]. Available: `http://www.boeing.com/commercial/noise/` (visited on 04/16/2014).

[15] *Luchtverkeersbesluit Schiphol*, 2011. [Online]. Available: `https://wetten.overheid.nl/BWBR0014330`.

[16] ICAO, *Annex 11 to the Convention on Civil Aviation: Air Traffic Services*, 2003.

[17] ——, *Doc 4444: Procedures for Air Navigation Services, Air Traffic Management*, 2001.

[18] "Guidelines for the application of the ECAC radar separation minima," EUROCONTROL, Tech. Rep., 1998.

[19] *Air Traffic Organisation Policy*, 2011.

[20] *UK AIC P064/2009*, Hounslow, UK, 2009.

[21] "Guidance on A380-800 Wake Vortex Aspects," ICAO, Tech. Rep., 2008.

[22] "RECAT-EU Edition 1.1," EUROCONTROL, Tech. Rep., 2015.

[23] FAA, *Fact Sheet - Wake RECAT*. [Online]. Available: `https://www.faa.gov/news/fact_sheets/news_story.cfm?newsId=18676` (visited on 06/02/2020).

[24] *Baangebruik Schiphol*, 2012. [Online]. Available: `http://www.lvnl.nl/nl/omgeving/baangebruik.html` (visited on 04/10/2012).

[25] BAA, *Heathrow Noise*, 2012. [Online]. Available: `http://www.heathrowairport.com/noise/` (visited on 04/10/2012).

[26] N. T. Ho and J.-P. B. Clarke, "Mitigating operational aircraft noise impact by leveraging on automation capability," in *AIAA Aviation Technology, Integration, and Operations Conference (ATIO)*, Los Angeles, CA, USA, 2001.

[27] L. Ren and J.-P. B. Clarke, "Flight-test evaluation of the Tool for Analysis of Separation and Throughput," *Journal of Aircraft*, vol. 45, no. 1, pp. 323–332, Jan. 2008.

[28] R. Mogford, J. Guttman, S. Morrow, and P. Kopardekar, "The complexity construct in air traffic control: A review and synthesis of the literature.," FAA, Washington, DC, USA, Tech. Rep. July, 1995.

[29] "D2 - The performance target," SESAR Consortium, Brussels, Belgium, Tech. Rep., 2006.

[30] P. M. A. de Jong, N. de Gelder, R. P. M. Verhoeven, F. J. L. Bussink, A. P. L. A. Marsman, and M. Mulder, "Aircraft noise and emission reduction through time and energy management during descent and approach," in *Proceedings of the Aircraft Noise and Emission Reduction Symposium*, Marseille, France, 2011.

[31] J. Shepley, "Near-term terminal area automation for arrival coordination," in *8th USA/Europe Air Traffic Management Research and Development Seminar*, Napa, CA, USA, 2009.

[32] A. Haraldsdottir, J. Scharl, J. King, and M. E. Berge, "Arrival management architecture and performance analysis with advanced automation and avionics capabilities," in *AIAA Aviation Technology, Integration, and Operations Conference (ATIO)*, Hilton Head, SC, USA, Sep. 2009.

[33] AviBit, *Optamos arrival management system*, Graz, Austria, 2012.

[34] M. Kupfer, T. J. Callantine, L. Martin, J. Mercer, and E. A. Palmer, "Controller support tools for schedule-based terminal-area operations," in *9th USA/Europe Air Traffic Management Research and Development Seminar*, Berlin, Germany, 2011.

[35] L. Ren, J.-P. B. Clarke, and N. T. Ho, "Achieving low approach noise without sacrificing capacity," in *IEEE/AIAA Digital Avionics Systems Conference*, AIAA, Piscataway, NJ, USA: IEEE, 2003, 1.E.3–1.1.

[36] J. Thipphavong and D. Mulfinger, "Design considerations for a new terminal area arrival scheduler," in *AIAA Aviation Technology, Integration, and Operations Conference (ATIO)*, Fort Worth, TX, USA, 2010.

[37] J. Bronsvoort, G. McDonald, M. Paglione, C. Garcia-Avello, I. Bayraktutar, and C. M. Young, "Impact of missing longitudinal aircraft intent on descent trajectory prediction," in *30th IEEE/AIAA Digital Avionics Systems Conference*, Seattle, WA, USA, Oct. 2011.

[38] "ATFCM users manual, Edition 16.0," EUROCONTROL, Brussels, Belgium, Tech. Rep., 2012.

[39] J. K. Klooster, K. D. Wichman, and O. F. Bleeker, "4D Trajectory and time-of-arrival control to enable continuous descent arrivals," in *AIAA Guidance, Navigation, and Control Conference*, Honolulu, HI, USA, Aug. 2008.

[40] J. K. Klooster and D. de Smedt, "Controlled time-of-arrival spacing analysis," in *9th USA/Europe Air Traffic Management Research and Development Seminar*, Berlin, Germany, Jun. 2011.

[41] M. R. C. Jackson, "CDA with RTA in a mixed environment," in *2009 IEEE/AIAA 28th Digital Avionics Systems Conference*, Orlando, FL, USA: Ieee, 2009, pp. 2.C.2–1–2.C.2–10.

[42] G. McDonald and J. Bronsvoort, "Concept of operations for Air traffic management by managing uncertainty through multiple metering points," in *Air Transport and Operations: Proceedings of the Third International Air Transport and Operations Symposium 2012*, Delft, The Netherlands: IOS Press, Jun. 2012.

[43] H. J. D. Reynolds, T. G. Reynolds, and R. J. Hansman, "Human factors implications of continuous descent approach procedures for noise abatement in air traffic control," in *6th USA/Europe Air Traffic Management Research and Development Seminar*, Baltimore, MD, USA, 2005.

[44] S. Mondoloni, S. Swierstra, and M. Paglione, "Assessing trajectory prediction performance - metrics definition," in *24th Digital Avionics Systems Conference*, vol. 1, Washington, DC, USA: IEEE, Oct. 2005.

[45] S. Mondoloni, M. Paglione, and S. M. Green, "Trajectory modelling accuracy for ATM decision support tools," in *International Congress of Aeronautical Sciences (ICAS)*, Toronto, Canada, Sep. 2002.

[46] T. Mueller, J. Sorensen, and G. Couluris, "Strategic aircraft trajectory prediction uncertainty and statistical sector traffic load modeling," in *AIAA Guidance, Navigation, and Control Conference and Exhibit*, Reston, VA, USA, Aug. 2002.

[47] G. Hunter, "Toward a standardized reference set of trajectory modeling errors," in *AIAA Modelling and Simulation Technologies Conference and Exhibit*, Providence, RI, USA, Aug. 2004.

[48] "White paper: Common TP structure and terminology in support of SESAR & NextGen," Eurocontrol/FAA Action Plan 16, Tech. Rep., 2010.

[49] R. A. Vivona, K. T. Cate, and S. M. Green, "Comparison of aircraft trajectory predictor capabilities and impacts on automation interoperability," in *11th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, Virginia Beach, VA, USA, Sep. 2011.

[50] R. A. Coppenbarger, "Climb trajectory prediction enhancement using airline flight-planning information," in *AIAA Guidance, Navigation, and Control Conference*, San Diego, CA, USA, 1999.

[51] T. J. Callantine and E. A. Palmer, "Controller advisory tools for efficient arrivals in dense traffic environments," in *9th AIAA Aviation Technology, Integration, and Operations Conference (ATIO)*, Hilton Head, SC, USA, Sep. 2009.

[52] R. E. Cole, S. M. Green, M. Jardin, B. E. Schwartz, and S. G. Benjamin, "Wind prediction accuracy for air traffic management decision support tools," in *3rd USA/Europe Air Traffic Management Research and Development Seminar*, Napoli, Italy, 2000.

[53] S. de Haan, "High resolution temperature and wind observations from commercial aircraft," in *International Symposium on Tropospheric Profiling*, Delft, The Netherlands, 2009.

[54] A. M. P. de Leege, M. Mulder, and M. M. van Paassen, "Novel method for wind estimation using Automatic Dependent Surveillance-Broadcast," *Journal of Guidance, Control, and Dynamics*, vol. 35, no. 2, pp. 648–653, Mar. 2012.

[55] S. de Haan and A. Stoffelen, "Assimilation of high-resolution Mode-S wind and temperature observations in a regional NWP model for nowcasting applications," *Weather and Forecasting*, vol. 27, no. 4, pp. 918–937, Aug. 2012.

[56] Barco N.V., *OSYRIS Trajectory Predictor Brochure*, 2011.

[57] D. Poles and A. Nuic, "Advanced aircraft performance modeling for ATM: Analysis of BADA model capabilities," in *29th Digital Avionics Systems Conference*, Salt Lake City, UT, USA, 2010.

[58] Skyguide, *ATM solutions factsheet: CALM*, 2012.

[59] H. N. Swenson, T. Hoang, S. Engelland, *et al.*, "Design and operational evaluation of the traffic management advisor at the Fort Worth Air Route Traffic Control Center," in *1st USA/Europe Air Traffic Management Research and Development Seminar*, Saclay, France, 1997.

[60] M. Porretta, M.-D. Dupuy, W. Schuster, A. Majumdar, and W. Ochieng, "Performance evaluation of a novel 4D trajectory prediction model for civil aircraft," *Journal of Navigation*, vol. 61, no. 03, Jun. 2008.

[61] A. M. P. de Leege, M. M. van Paassen, and M. Mulder, "A machine learning approach to trajectory prediction," in *AIAA Guidance, Navigation, and Control (GNC) Conference*, Reston, VA, USA: American Institute of Aeronautics and Astronautics, Aug. 2013.

[62] ICAO, *Doc 9613: Performance-Based Navigation (PBN) manual*, 2008.

[63] J. Bronsvoort, G. McDonald, R. Potts, and E. Gutt, "Enhanced descent wind forecast for aircraft," in *9th USA/Europe Air Traffic Management Research and Development Seminar*, Berlin, Germany, 2011.

[64] A. van der Eijk, M. Mulder, M. M. van Paassen, and A. C. in 't Veld, "Assisting air traffic control in planning and monitoring continuous descent approach procedures," in *AIAA Guidance, Navigation, and Control Conference*, ser. AIAA Guidance, Navigation, and Control Conference, Chicago, IL, USA: American Institute of Aeronautics and Astronautics (AIAA), Aug. 2009.

[65] NASA, *Traffic Management Advisor*, 2011. [Online]. Available: `http://www.aviationsystemsdivision.arc.nasa.gov/research/foundations/tma.shtml` (visited on 02/12/2012).

[66] S. S. Potter and D. D. Woods, "Event-driven timeline displays: Beyond message lists in human-intelligent system interaction," in *International Conference on Systems, Man, and Cybernetics*, Columbus, OH, USA: IEEE, Oct. 1991.

[67] L. Bernard-Peyre, “A smoother rhythm,” *Air Traffic Technology International*, pp. 94–97, 2010.

[68] Barco N.V., *OSYRIS Queue Management brochure*, 2011.

[69] W. Nedell, H. Erzberger, and F. Neuman, “The Traffic Management Advisor,” in *Proceedings of the American Control Conference*, San Jose, CA, USA: IEEE, May 1990, pp. 514–520.

[70] D. Chang and K. V. Nesbitt, “Identifying commonly-used gestalt principles as a design framework for multi-sensory displays,” in *International Conference on Systems Man and Cybernetics*, Taipei, Taiwan, 2006.

[71] E. G. Knapen, “ERAT Arlanda RTS controller handbook,” NLR, Amsterdam, The Netherlands, Tech. Rep., 2010.

[72] M. Tielrooij, A. C. in ’t Veld, M. M. van Paassen, and M. Mulder, “Development of a time-space diagram to assist ATC in monitoring continuous descent approaches,” in *Air Traffic Control*, M. Mulder, Ed., Rijeka, Croatia: InTech, 2010, pp. 135–147. [Online]. Available: `http://www.intechopen.com/books/air-traffic-control/development-of-a-time-space-diagram-to-assist-air-traffic-controllers-in-monitoring-continuous-desce`.

[73] E. van Dijk, M. Mulder, M. M. van Paassen, and M. Roerdink, “An Interface for Inbound Traffic Management by Air Traffic Control,” in *AIAA Guidance, Navigation, and Control Conference*, Chicago, IL, USA, Aug. 2009.

[74] R. E. Klomp, M. Mulder, and M. M. van Paassen, “Re-design of an inbound planning interface for air traffic control,” in *AIAA Guidance, Navigation, and Control Conference*, Portland, OR, USA: American Institute of Aeronautics and Astronautics (AIAA), 2011.

[75] A. M. P. de Leege, A. C. in ’t Veld, M. Mulder, and M. M. van Paassen, “A time-space diagram as controller support tool for closed path continuous descent operations,” in *AIAA Guidance, Navigation, and Control Conference*, Portland, OR, USA: American Institute of Aeronautics and Astronautics (AIAA), 2011.

[76] R. A. Coppenbarger, G. Dyer, M. Hayashi, R. C. Lanier, L. Stell, and D. N. Sweet, “Development and testing of automation for efficient arrivals in constrained airspace,” in *International Congress of Aeronautical Sciences (ICAS)*, Nice, France, 2010.

[77] M. Tielrooij, “Assisting ATC in supervising continuous descent approach procedures,” MSc. Thesis, Delft University of Technology, The Netherlands, 2008.

[78] M. Whiteley, “PHARE advanced tools tactical load smoother final report,” EUROCONTROL, Brussels, Belgium, Tech. Rep. November, 1999.

[79] Barco N.V., *Air Traffic Management software: Queue Management*, 2012.

[80] J. Nibourg, H. Hesselink, and Y. A. J. R. van de Vijver, "Environment-Aware Runway Allocation Advice System," in *26th International Congress of the Aeronautical Sciences*, Anchorage, AK, USA, 2008.

[81] H. Hesselink and J. Nibourg, "Probabilistic 2-day forecast of runway use," in *9th USA/Europe Air Traffic Management Research and Development Seminar*, Berlin, Germany, 2011.

[82] E. G. Knapen, "Applicability of arrival management systems for schiphol airport," NLR, Amsterdam, The Netherlands, Tech. Rep., 2011.

[83] D. Schaefer, A. Gizdavu, and D. Nicholls, "The display of uncertainty information on the working position," in *23rd IEEE/AIAA Digital Avionics Systems Conference*, vol. 1, Washington, DC, USA, Oct. 2004.

[84] P. Whysall, "Future Area Control Tools Support (FACTS)," in *2nd USA/Europe Air Traffic Management Research and Development Seminar*, Orlando, FL, USA, Dec. 1998.

[85] K. J. Vicente and J. Rasmussen, "Ecological interface design: theoretical foundations," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 22, no. 4, pp. 589–606, 1992.

[86] J. Rasmussen, "A taxonomy for analysis of cognitive work," in *IEEE Conference on Human Factors and Power Plants*, Monterey, CA, USA: IEEE, Jun. 1992.

[87] A. Vanwelsenaere, J. Ellerbroek, J. M. Hoekstra, and E. Westerveld, "Analysis on the impact of pop-up flight occurrence when extending the arrival management horizon," in *12th USA/Europe Air Traffic Management Research and Development Seminar*, Seattle, WA, USA, Jun. 2017.

[88] C. E. Billings, *Aviation automation, the search for the human-centered approach*. Malwah, NJ: Lawrence Erlbaum Associates Inc., 1997.

[89] E. Oprins, D. Zwaaf, F. Eriksson, R. Roe, and K. van de Werwe, "Impact of future time-based operations on situation awareness of air traffic controllers," in *8th USA/Europe Air Traffic Management Research and Development Seminar*, Napa, CA, USA, 2009.

[90] K. Christoffersen and D. Woods, "How to make automated systems team players," in *Advances in Human Performance and Cognitive Engineering Research Vol. 2*, 2002, E. Salas, Ed., Emerald Group Publishing Limited, 2002, pp. 1–12.

[91] D. Nicholls, P. Marti, P. Battino, S. Pozzi, and V. Barsotti, "CARE innovative action presenting uncertainty to controllers & pilots," EUROCONTROL, Brussels, Belgium, Tech. Rep. December, 2004, pp. 1–53.

[92] C. Wanke, S. Mulgund, D. Greenbaum, and L. Song, "Modeling traffic prediction uncertainty for traffic management decision support," in *AIAA Guidance, Navigation, and Control Conference and Exhibit*, Reston, VA, USA: AIAA, Aug. 2004.

Chapter 3

Predicting ETA uncertainty from flight information

The previous chapter proposed the introduction of knowledge on the uncertainty in arrival management to overcome its limiting effects on the AMAN working horizon. Such an approach will require a means of determining uncertainty. This chapter explores the use of empirical information on errors in flight update messages (FUMs) from EUROCONTROL's Network Manager.
This chapter is based on a paper written in 2014. The references to the current time in that publication designate the time of writing. The actual present (2022) will be explicitly indicated as such. At the time of writing, the FUM was the only information readily available at Air Traffic Control the Netherlands. Newer message formats and contents had been developed by EUROCONTROL at that time and are nowadays (2022) used by LVNL. However, the method presented in this paper could also apply to these more modern equivalents.

This chapter is based on the following publication:

Paper Title	Predicting arrival time uncertainty from actual flight information
Authors	M. Tielrooij, C. Borst, M.M. van Paassen and, M. Mulder
Published in	11th USA/Europe Air Traffic Management Research and Development Seminar, June 2015, Lissabon, Portugal (pp. 577-586)

.

3.1 Introduction

Many Air Navigation Service Providers (ANSPs) nowadays use Arrival Managers (AMANs) to plan the arrival times of inbound aircraft and thereby balance the demand to the available capacity. These systems support the sequence manager in deciding how to plan arrival times of aircraft. Such planning is required when the intervals between predicted landing times of consecutive flights become smaller than the required spacing.

When aircraft are assumed to fly the optimal trajectory from an Airspace User's (AU) perspective, any deviations in the planned 4D path from these trajectories imply a reduction of efficiency. Air Traffic Control (ATC) should aim to minimize such deviations. If changes are required, however, earlier decisions allow for smaller trajectory changes, increasing flight efficiency. For example, a modest speed increase over a longer flight time is often much more fuel-efficient than a substantial speed increase over a shorter time while achieving the same difference in time.

Currently, the horizons at which AMANs are used to monitor and influence traffic are limited by three factors to a horizon of typically 20 to 30 minutes (or 150-200 NM) [1]:

1. The availability of information on the predicted arrival time of aircraft, due to the limit of radar surveillance, for example,

2. The authority to influence the flight's trajectory. For example, the boundaries of Flight Information Regions (FIRs),

3. The reliability of the predicted arrival times.

Future operational concepts, such as those that are proposed in Single European Sky ATM Research (SESAR) and Next Generation Air Traffic Management System (NextGen), foresee an increase in the planning horizon to increase long-term efficiency and predictability of operations. Depending on the concept, future horizons are expected to be at 200-500 NM, or about two hours of flying time [2], [3]. Achieving this increase in a horizon requires a reduction of the three limitations on the current AMAN.

The first two of the constraints listed above are being addressed in current developments: System Wide Information Management (SWIM) should support continuous sharing of all relevant information concerning a flight between all actors involved [4], [5]. And, through SWIM, different ANSPs are expected to be able to share their requirements on a trajectory (such as an arrival time planned by AMAN) and their ability to achieve those requirements. While these developments resolve the first two limitations on the planning horizon listed above, the third problem, prediction uncertainty, is expected to reduce but is unlikely to disappear altogether. If uncertainty is not eliminated, increasing the AMAN horizon will require ways to perform arrival planning in the presence of uncertainty.

Any approach that makes uncertainty an integral part of the decision-making process will require an estimate of the uncertainty. This chapter will propose a method to predict the uncertainty on an Estimated Time of Arrival (ETA) based on currently available information on flight progress.

The next section discusses current methods to predict arrival time uncertainty. Sections 3.3 and 3.4 will then analyse the prediction errors in EUROCONTROL's Flight Update Messages (FUMs) that ANSPs currently use to report flight progress. This analysis will use flight data for Amsterdam Airport Schiphol in 2013 and 2014. Based on this analysis and the different modelling techniques, Section 3.5 will then describe a method to predict uncertainty by estimating distributions using historical data. The method is validated in Section 3.6. Finally, Section 3.7 discusses the results of this validation and the applicability of the method to other airports and other data sources.

Throughout this chapter, the term uncertainty will be used to describe the expectation of the difference between a prediction of an arrival time and the actual arrival time. This uncertainty, therefore, includes both bias and noise. The resulting distributions from the proposed method ultimately provide an indication of both.

3.2 A review on modelling arrival time uncertainty

Past research on prediction uncertainty addresses both strategic flow management and tactical Medium Term Conflict Detection (MTCD). This body of research provides the different causes of prediction error and their effects on the uncertainty with respect to the prediction horizon.

3.2.1 Prediction uncertainty

Current operational prediction capabilities depend strongly on the phase of flight. Typical standard deviations for airborne flights are 30 seconds at 20 minutes before landing when airborne (e.g., Flight Management System (FMS) predictions as analysed by Bronsvoort [6]). These standard deviations increase to 15 minutes when the aircraft is still on the ground (e.g., the departure accuracies of several airports in the US found by Mueller and Chatterji [7]).

Prior research on prediction of uncertainty in Air Traffic Management (ATM) typically focuses on either short horizons (i.e., 20 minutes) for tactical tools, or very long horizons (i.e., multiple hours) for flow management and strategic purposes [8]–[11]. In the future ATM concepts, AMAN concepts will require an estimate of the uncertainty at the horizon limit of two hours falling between those two scopes.

Most—if not all—aircraft will be airborne at short horizons, and their current, up-to-date states are available through surveillance techniques such as radar and Automatic Dependent Surveillance Broadcast (ADS-B). To indicate accuracy, a study of airborne FMS prediction capabilities shows a typical standard deviation of 60 seconds at 30 minutes before arrival and 120 seconds at 60 minutes before arrival, with uncertainty rapidly rising beyond that horizon [6]. Ground-based prediction systems are likely to have less information on the actual state of the aircraft and therefore increase the potential for errors [12]. On the other hand, the airborne system may be unaware of future ATC decisions that may affect the aircraft's trajectory.

The proximity of departure airports makes flight status—e.g. whether it is airborne or not—is an important factor in determining prediction uncertainty. The effects of the flight status on prediction error have been demonstrated in research by Solveling [13] and Tobaruela [11]. Many of the busiest connections of European airports are within a two-hour flight horizon. For example, the 20 busiest connections to Amsterdam are within this horizon (Figure 1.2) [14]. When the prediction horizon is longer than the flight time, the disturbances and uncertainties associated with taxiing, boarding, and the previous rotation of the aircraft are all added to the set of disturbances.

Since a future system should provide support in continuous decision-making during the tactical phase, the uncertainty associated with an ETA needs to be updated continuously as well. These updates require an approach that can be executed online during operation [13]. Hence, any method of predicting uncertainty has to support fast calculation and model the different uncertainties in the different phases of a flight.

3.2.2 Predicting uncertainty

To predict the uncertainty for a particular flight, the iFACTS and CARE projects employed methods based on the propagation of the uncertainty of the input components [8], [9]. Stepwise propagation allows for a detailed and fast prediction over the tactical horizon of approximately 20 minutes. To use this approach in the proposed AMAN context, adding sources of error at much longer time horizons requires knowledge of the uncertainty associated with those errors. The inclusion of more components implies more complicated algorithms and a more comprehensive understanding of the behaviour of all contributing components. At longer horizons, this could become impracticable due to the increasing complexity of that behaviour and, subsequently, the computative load required in combining those models.

To support research into the effects of airspace changes, Wanke et al. developed a Monte-Carlo approach that uses empirical information on the outcomes of predictions to model uncertainty [15]. The authors suggest the possible value of such an approach in tactical operation, given the availability of sufficient computing power. However, the method is currently unsuitable for real-time applications.

Both applications calculate the uncertainty for the 3D path and the flight time. Since AMAN aims to form a plan based on the arrival time, only the flight time is important. At the same time, ANSPs nowadays receive regular updates on the ETA for flights on the ground and in the air. Future concepts, such as in SESAR, foresee the expansion of this capability with the advent of SWIM and the Business Trajectory concept. The focus on arrival time only and the availability of ETA estimates provide a means to develop a method to predict uncertainty.

3.2.3 Using empirical information

Tobaruella et al. [11] use an experience-based approach to predict the number of flights in a sector given predictor uncertainty. This method uses historical error patterns to estimate errors on ETFMS Flight Data (EFD) messages provided by the EUROCONTROL Network Manager (NM). Such a method directly relates prediction error to the actual information available on the flight and the properties of the prediction in particular. This method allows for an uncertainty predictor sensitive to all relevant aspects and fast enough for real-time application. If an error distribution can be captured in a small set of parameters, a look-up table with those parameters, based on information in the arrival time prediction, may provide a rapid way of predicting error distributions.

Our study focuses on arrival planning for Amsterdam. Since EFD data are only recently made available, Air Traffic Control the Netherlands (LVNL) currently does not have an extensive recording of those messages, precluding any prediction error analysis. However, the critical information available in EFD messages is also available in FUMs. These messages are currently provided by the NM over the Aeronautical Fixed Telecommunication Network (AFTN) and are, therefore, available. Using this information, a technique similar to that of Tobaruella is also applicable in current operations.

3.3 Flight Update Messages

The FUM is a message that provides downstream ANSPs with the progress of a flight [16]. As such, the FUM contains an estimate of the ETA of a flight at the destination airport. Recording these ETAs and the actual arrival times provides a set of prediction errors. The distribution of these errors indicates the prediction uncertainty in the FUMs. This section describes the FUM, its generation, and when the messages are communicated.

3.3.1 Content of the message

The messages contain:

- Origin and destination airport;
- Estimated Off-Block Time (EOBT);
- Estimated Landing Time (ELDT);
- Aircraft type;
- Expected Standard Terminal Arrival Route (STAR);
- Estimated level at STAR waypoint;
- Estimated time at STAR waypoint;
- Flight status, and
- CDM Status (when available).

3.3.2 Flight status

The flight status indicates the flight's progress from filing the flight plan until the aircraft has arrived. As each flight phase has its own set of potential disturbances, the flight status is an important element in the uncertainty. The possible flight statuses in a FUM are [16]:

- FI - Filed: A flight plan has been filed, but no other information is available.
- SI - Slot Issued: A slot has been allocated by the EUROCONTROL Network Manager to manage congestion.
- TA - Tactical Activated: The Enhanced Tactical Flow Management System (ETFMS) assumes the flight to be airborne. To maintain a correct traffic picture, NM assumes the flight to be airborne when its slot, or Calculated Time of Take-off (CTOT), has passed.
- AA - ATC Activated: The flight is confirmed to be airborne by the first ANSP to have the flight under its control.
- CA - Cancelled: The operator has cancelled the flight.
- SU - Suspended: The ETFMS is no longer predicting the flight's progress. Any following ETA is therefore no longer valid.
- TE - Terminated: This message is provided after confirmation of arrival has been received from the destination airport or 20 minutes after the ELDT. In the first case, the destination airport reports the landing time.

3.3.3 Prediction

The prediction is based on the flight plan filed by the AU, the predicted wind and the EOBT. From the EOBT, a fixed taxi duration is assumed based on the departure airport. From the take-off time, a prediction is made based on the aircraft model, the filed route, and the predicted winds.

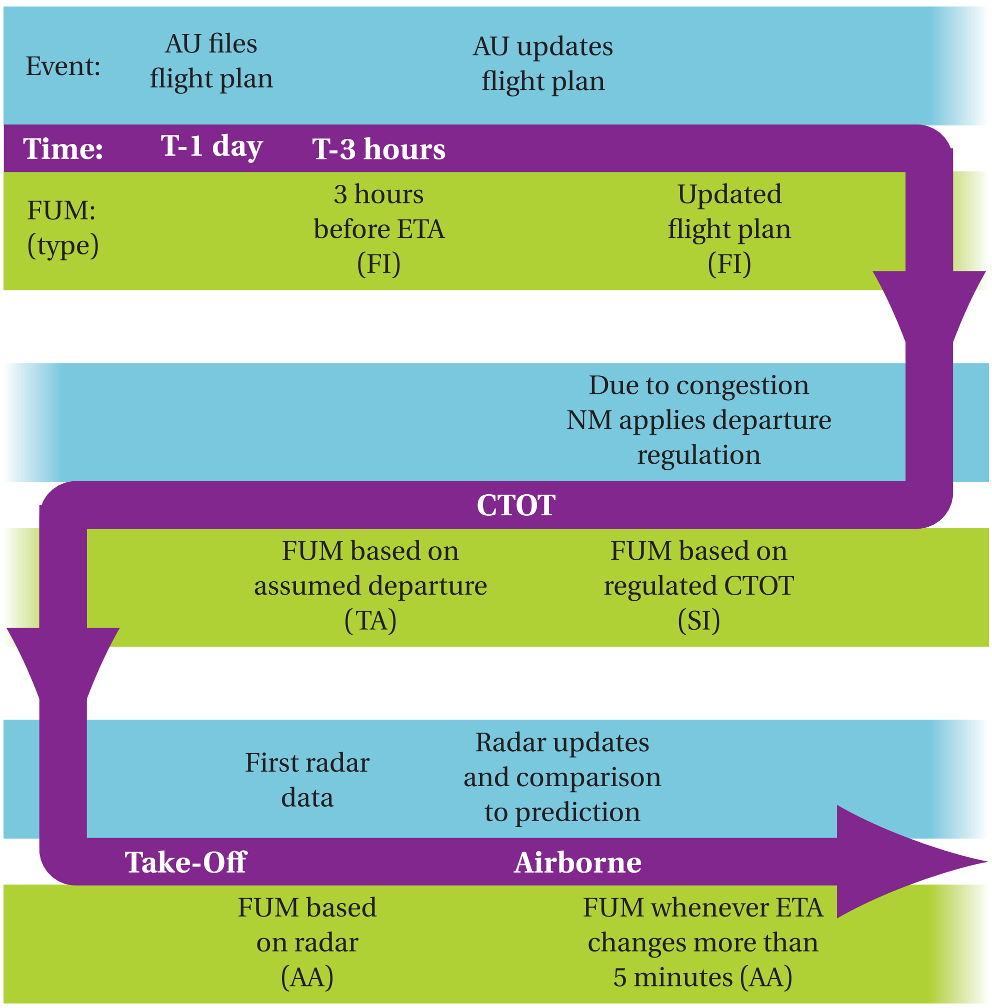


Figure 3.1: FUM generation and update process.

Figure 3.1 provides an overview of when a FUM update is provided. The NM generates a prediction at regular intervals and significant changes to the flight. These changes mainly concern changes in the flight status or changes in Airport Collaborative Decision Making (A-CDM) status. Subsequently, the predicted arrival time is compared to the previously published arrival time. NM publishes a new ETA if these differ more than five minutes or if a status change has occurred.

The fact that an update is only provided when the prediction error is greater than five minutes implies that a prediction has a potential undetected error of five minutes and an equally large noise in the ETA. Since FUMs are only used by ANSPs for initial planning, this error is acceptable for planning purposes. However, for precise planning such as required in AMAN, this accuracy is insufficient as landing intervals are typically around one minute [17].

3.4 Analysis of FUM estimates

This study uses all data for 2013 and 2014 for Amsterdam Airport Schiphol. A full log of all received AFTN messages has been filtered for relevant messages on inbound flights. These messages were correlated to each flight based on the callsign, the flight date, departure airport, and arrival airport. The correlated messages were stored in a database. The messages were then processed to generate histograms of errors from which distributions could be estimated. Figure 3.2 visualises the processing of the messages.

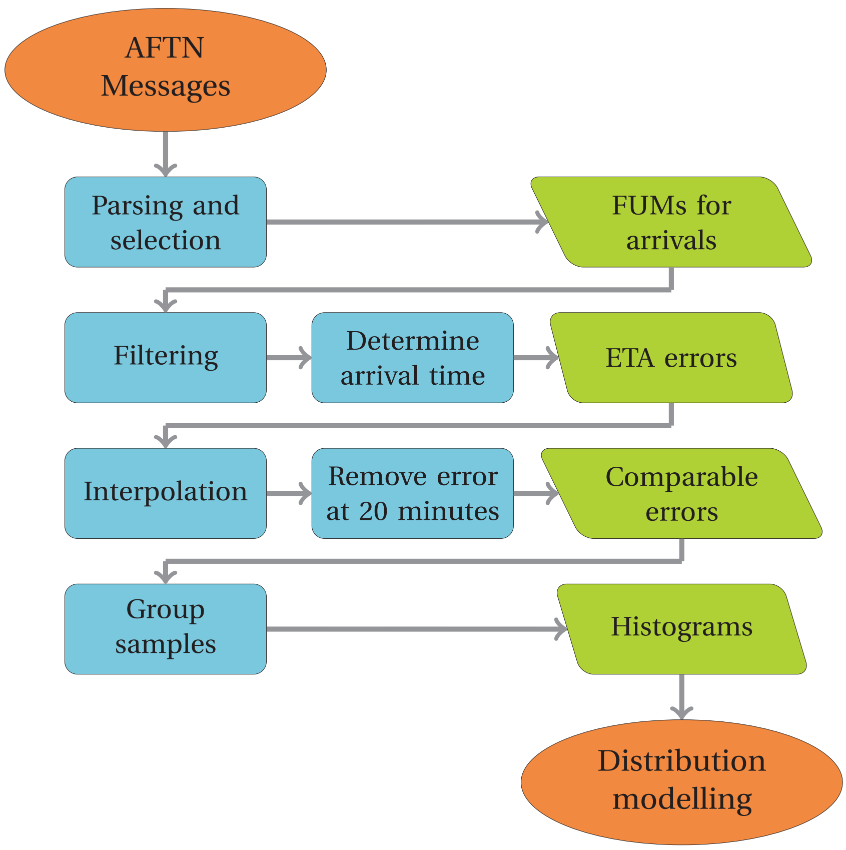


Figure 3.2: Process to develop error distributions from original AFTN messages.

Several steps required results from later analyses. The next sections follow the order in which the analyses were originally developed. The final process followed these steps in order:

1. Filtering messages ensures all required information is available in each message (Section 3.4.1).
2. The last message for each flight provides the reference landing time (Section 3.4.2).
3. Since the messages do not arrive at fixed times, interpolation of the error ETAs ensures that messages of different flights can be compared (Section 3.4.4).
4. To avoid the inclusion of prediction error due to the assumed route in the arrival airspace, the prediction error is corrected for the error before entering that airspace (Secton 3.4.2).
5. Grouping errors with identical parameters (e.g., type of messages, remaining time to fly) forms histograms of the prediction accuracy (Section 3.4.3).
6. Finally, these histograms are input to the estimation of distributions.

3.4.1 Filtering

Initial filtering of the messages applied the following rules to reduce erroneous or ambiguous data:

- The message format matches the standardised formats.
- Messages should contain all elements necessary for correlation with a particular flight (callsign, departure date, origin, destination).
- A flight plan message— which contains essential information on the flight— must have been received before any other messages on that flight are accepted.
- A terminator message—which provides the actual landing time —has been received.
- Flights were not cancelled or diverted.
- Flights have no errors of greater than 12 hours. Close inspection of these flights showed that these are mostly incorrectly correlated flights of the following day that use the same callsign.
- The flight did not occur on one of five days with excessive disruption (due to wind, fog, or industrial action). The planning process deviates a lot from the normal operation to accommodate the uncertainty on such days. These days are identified by comparing the mean and standard deviation of the arrival delay to the scheduled arrival time and correlating the results to weather reports and posterior reports on disruptive events published by the NM [18], [19].

3.4.2 Determining the reference time

The Estimated Landing Time (ELDT) that was reported in the last message for a flight is assumed to be the actual landing time for that flight. In all cases, these messages were termination messages (TE). For the termination message, the term *estimated landing time* is misleading as this is the *actual* arrival time.

The prediction by the NM assumes a particular approach route to the Initial Approach Fix (IAF) and is unaware of the runway in use. Any difference between the assumed arrival route and the actual route will result in a prediction error. When approaching Amsterdam, aircraft are flying published routes only during the night. During the day, Air Traffic Controllers (ATCOs) will direct the traffic using vectors. One of the sources of information used in assigning vectors is the information from the AMAN. Therefore, the resulting changes in estimated landing times are not meaningful in the context of arrival management.

The selected runway and arrival route are a consequence of the AMAN prediction. When modelling the uncertainty—and in particular, how it changes during the evolution of the flight—the error in predicted arrival time due to the uncertainty in the arrival route is irrelevant. To avoid focussing on the approach phase, the analysis assumes that the error is zero at 20 minutes before arrival (i.e., before the start of the approach routes) to avoid this error. This approach assumes that the estimate is not updated beforehand based on updated knowledge of the approach route.

The correction is calculated by taking the error for a flight at 20 minutes and subtracting that error from all the flight's predictions. As a consequence, however, flights without a valid interpolation at 20 minutes before arrival have been discarded. Similarly, any messages received less than 20 minutes before arrival are removed from the analysis.

3.4.3 Grouping comparable messages

The total AFTN dataset contained messages for 446,147 flights that were successfully parsed. The airport has handled approximately 442,000 flights in that period [20], [21]. The difference is likely to be explained by duplicated flights in the dataset. After applying the filtering criteria, 398,328 flights remained, providing 1,628,072 FUMs. After correcting the data for the error at 20 minutes, 372,998 flights originating from 742 unique airports remained. The analysis includes 1,330,749 for these flights.

Table 3.1 lists the number of flights with the same number of messages. The highest value in each category is highlighted. Figure 3.3) shows the number of messages per flight of any type. The number of messages per flight varied with a median of 3.0, with only 5850 flights for which a single message was received.

Figure 3.4 provides the number of messages of each type per flight for the most important message types. Figure 3.4(a) shows that the flight plan message is only updated occasionally (FI). If a flight is assigned a departure slot (SI), the slot is likely to be updated as shown in Figure 3.4(b). The same holds for flights that are assumed to be airborne (TA). Finally, when flights are airborne, a few flights trigger an update from the initial airborne message (AA).

Table 3.1: Number of flights per number of messages.

Type	Number of messages					
	0	1	2	3	4	5+
FI	118515	**214155**	33249	4743	1392	926
SI	**288635**	46321	22317	8807	3884	3016
TA	**198303**	91946	53769	18903	5404	4655
AA	0	**262892**	75374	24185	8295	2234
Any	0	5850	**118812**	97198	65577	85543

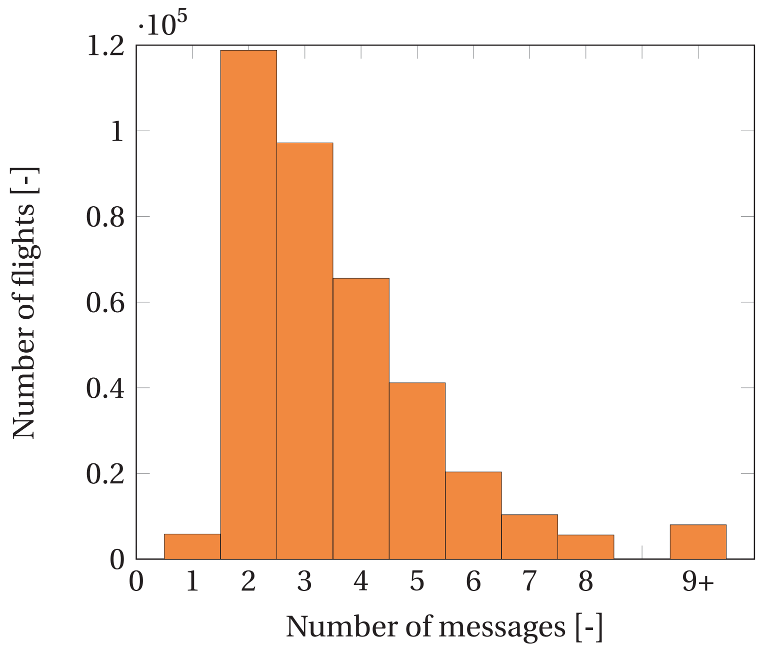
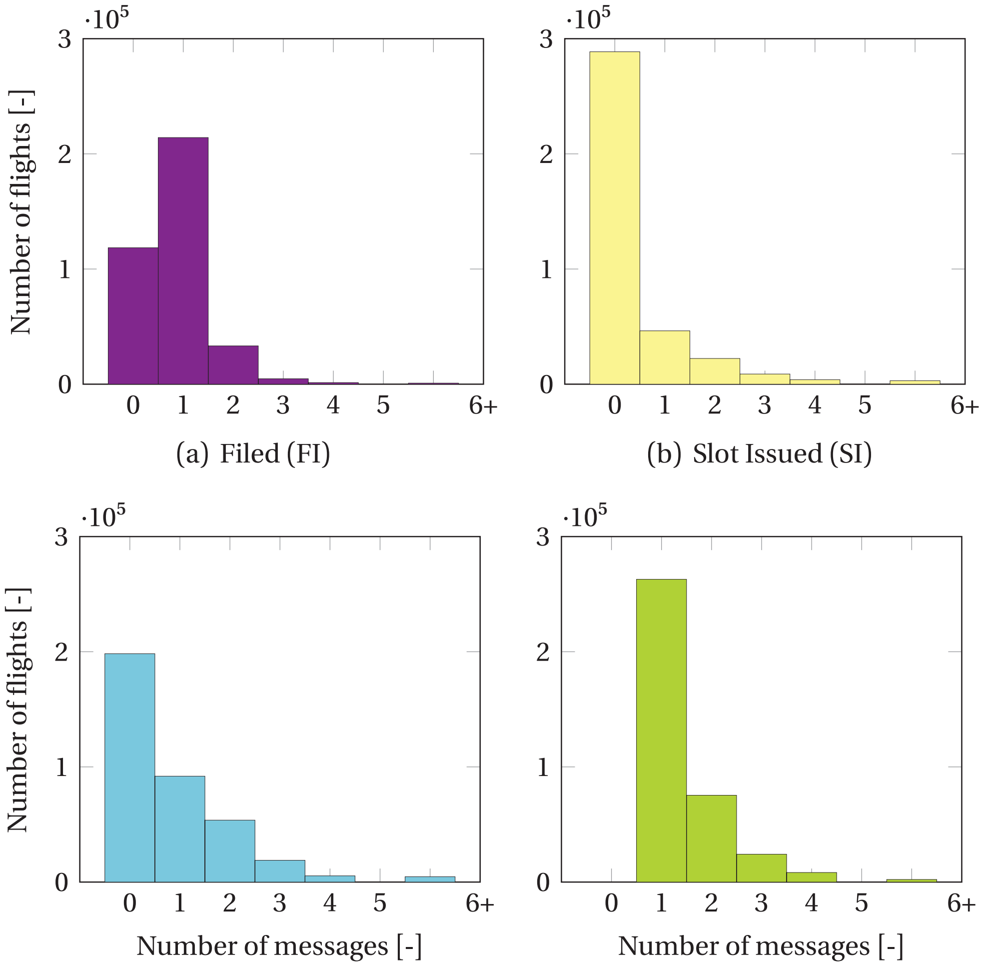


Figure 3.3: Distribution of the number of messages per flight.

Grouping comparable messages based on their properties enables the determination of the error distribution based on those properties. The following properties for a message are available in the available data and have been considered:



(a) Filed (FI)

(b) Slot Issued (SI)

(c) ETFMS Activated (TA)

(d) ATC Activated (AA)

Figure 3.4: Number of messages for each flight per message type.

- Flight callsign: Each flight may have particular objectives and punctuality requirements. However, most callsigns are only operated once a day and would provide very few samples per group. Furthermore, callsigns are not static and may change per day.

- Flight operator: Similar to the callsign, punctuality requirements may depend on the operator's business model. Again, some operators only have one flight per day, generating a small sample set.

- Flight status: As explained in Section 3.3.2, the flight status is likely to be a key factor in the prediction uncertainty.

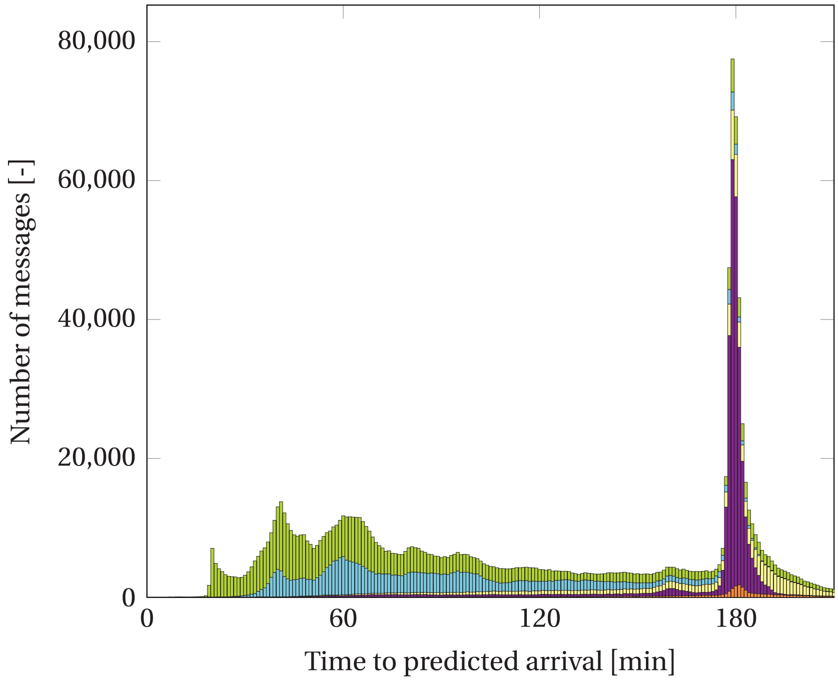
- Origin airport: Variation in airport operating procedures may influence punctuality. More importantly, the origin airport determines the flight's length and, therefore, whether an aircraft may still be on the ground within the horizon generating the associated additional uncertainty. However, this would be a large subset for the most frequent connections only. Also, the flight status indicator provides similar information on the flight's progress.

- Date of flight: The date of the flight may account for seasonal disruptive weather effects, such as de-icing during winter, for example.
- Planned date and time of arrival: The arrival date provides a similar indicator as the departure date. The arrival time may show daily effects due to congestion of airports and airspaces.
- Aircraft type: Different aircraft may be more flexible in achieving punctuality. Similarly, modern flight systems may provide more support for punctuality. But, again, only a few types generate most of the traffic. Other types would, therefore, form small samples.
- Remaining time before arrival: As already indicated, the prediction horizon is the most important parameter in prediction uncertainty.

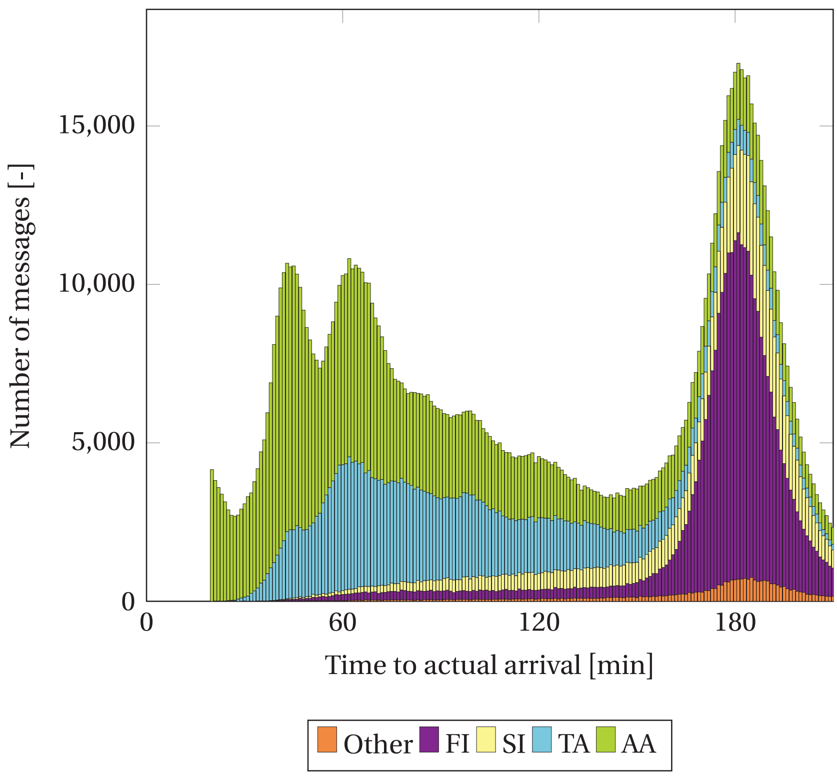
Selection of the criteria and the width of intervals for those criteria (bin width) requires a compromise between the detailing of the model and data availability. Based on an initial analysis of the available information, the following grouping criteria are selected: Horizon (in steps of ten minutes), flight status, and arrival time of day (in bins of three hours).

3.4.4 Interpolating errors

Figure 3.5 shows a distribution of the messages over the (predicted and actual) remaining time to arrival. In Figure 3.6 the number of messages have been separated by type. The graphs show how, closer to arrival, more and more airborne reports (AA) are provided and fewer messages with a flight status on the ground (FI or SI).

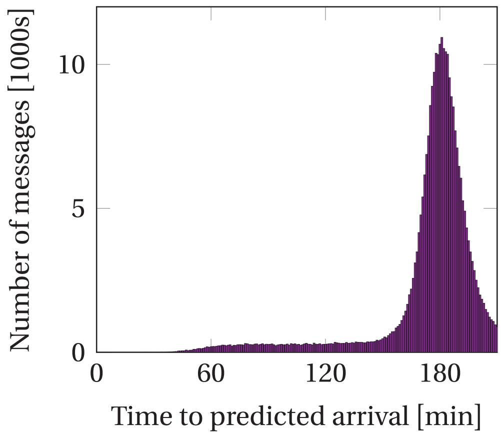


(a) Messagecount versus predicted time to fly.

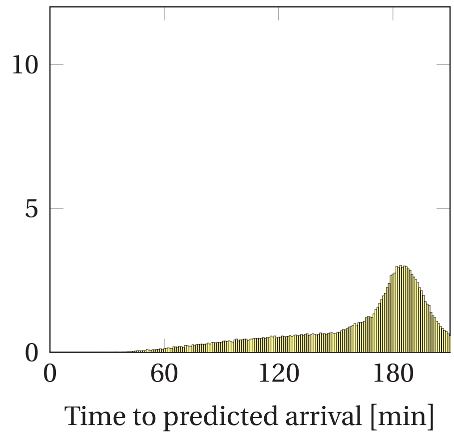


(b) Messagecount versus actual time to fly.

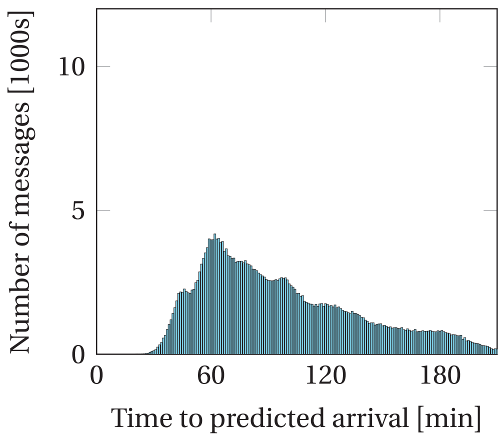
Figure 3.5: Distribution of messages and different flight statuses over the prediction horizon. The graphs show three clear peaks at 45, 65 and 180 minutes.



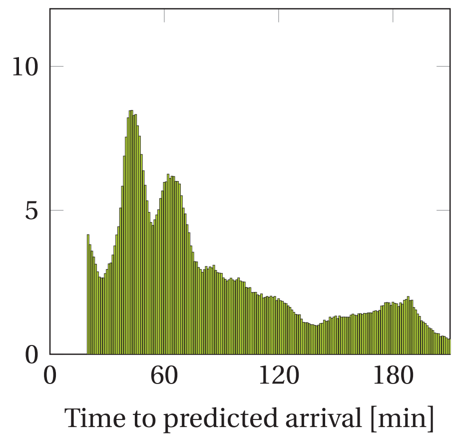
(a) Filed



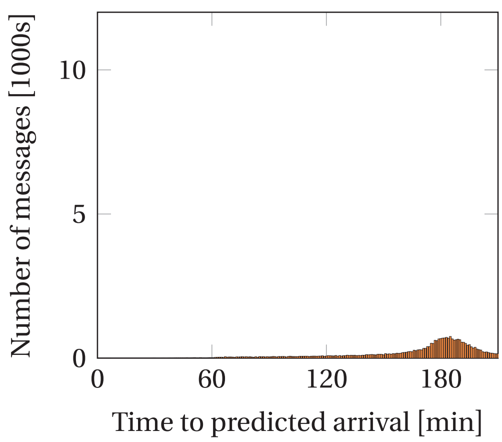
(b) Slot Issued



(c) ETFMS Activated



(d) ATC Activated

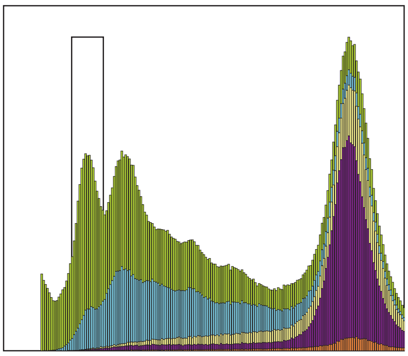


(e) Other

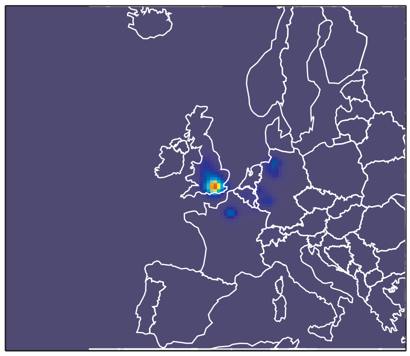
Figure 3.6: Message count by horizon per type.

The graphs show a peak in updates at three hours before the expected arrival, representing the publishing of the flight plan by NM 3 hours before the expected landing time. This peak is much broader when plotted against the actual time of arrival. While the peak has the same cause, the prediction uncertainty at that horizon—when the flight plans are published—causes considerable spread.

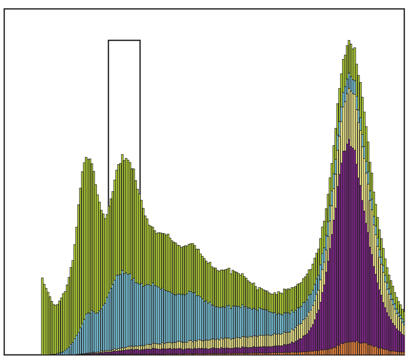
Figure 3.5(b) furthermore shows two noticeable peaks at around 45 and 65 minutes before arrival. In Figure 3.6, these peaks are shown to consist of AA-messages at 45 minutes, and TA and AA-messages at 65 minutes (See Subfigures 3.6(c) and 3.6(d)). Figure 3.7 shows that these relate to two groups of airports at a similar range. The first peak consists of aircraft departing from one of the London airports. The second peak contains flights from Edinburgh, Copenhagen, Berlin, Munich, and Zurich. All of which have a similar flying time to Amsterdam.



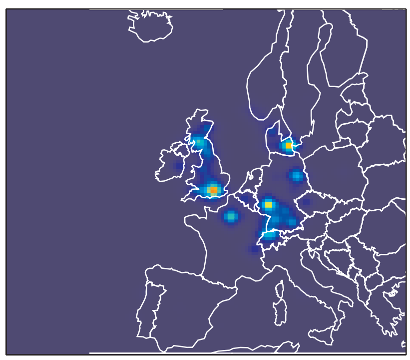
(a) First peak in message count



(b) Origin of flights in first peak



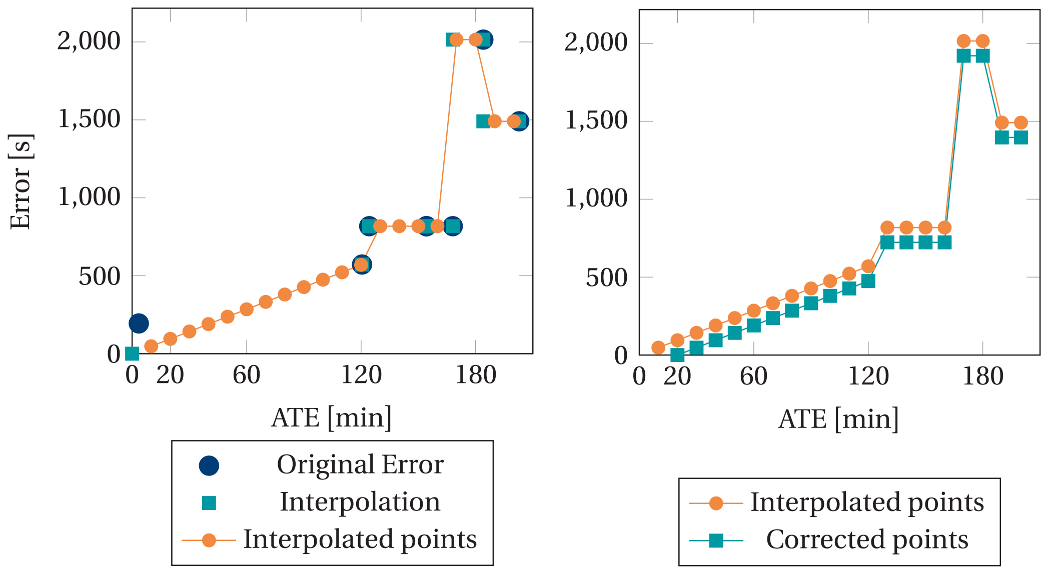
(c) Second peak in message count



(d) Origin of flights in second peak

Figure 3.7: Relation between peaks in number of messages and origin of flights.

The histograms indicate that the number of messages varies over the prediction horizon. Since the ETFMS only provides an update when the ETA deviates more than five minutes from the expected plan, accurately predicted flights are under-reported. Determining the accuracy based on the messages at a specific horizon would overestimate the uncertainty, especially at horizons close to arrival. To better represent the more accurate flights, the messages need interpolation.



(a) Interpolation based on the message type (b) Correction of the final prediction error

Figure 3.8: Error interpolation process and correction of the error at 20 minutes.

When aircraft are airborne, it is likely that the error gradually reduces, as less deviation is possible in the shorter remaining flight time. In this work, the errors for reports of airborne flights are assumed to be decreasing over time and therefore, linearly interpolated. This approach implicitly assumes that the error decreases linearly with horizon, which will be addressed in the discussion.

A flight status change is a discrete event; interpolating between such changes is not straightforward. Delaying events on the ground are more likely to be discrete, such as a report of a missing passenger. Errors for flight messages from flights that are still on the ground are assumed to be constant up to the following report.

Figure 3.8 shows the set of received predictions from an actual flight and the application of the interpolation process:

1. Any prediction less than 10 minutes from the actual arrival time is discarded as the remaining flight time is smaller than the interpolation interval.

2. The calculated errors are propagated forward to the error from the next message. The assumed error is constant for FI and SI messages as flights are on the ground. For TA and AA, the error changes linearly to the next value.

3. The propagated errors are interpolated at 10 minute intervals.

4. The interpolated error at 20 minutes is subtracted from all points. This subtraction discards the last interpolation point at 10 minutes as it is no longer meaningful.

3.4.5 Prediction accuracy versus time to fly

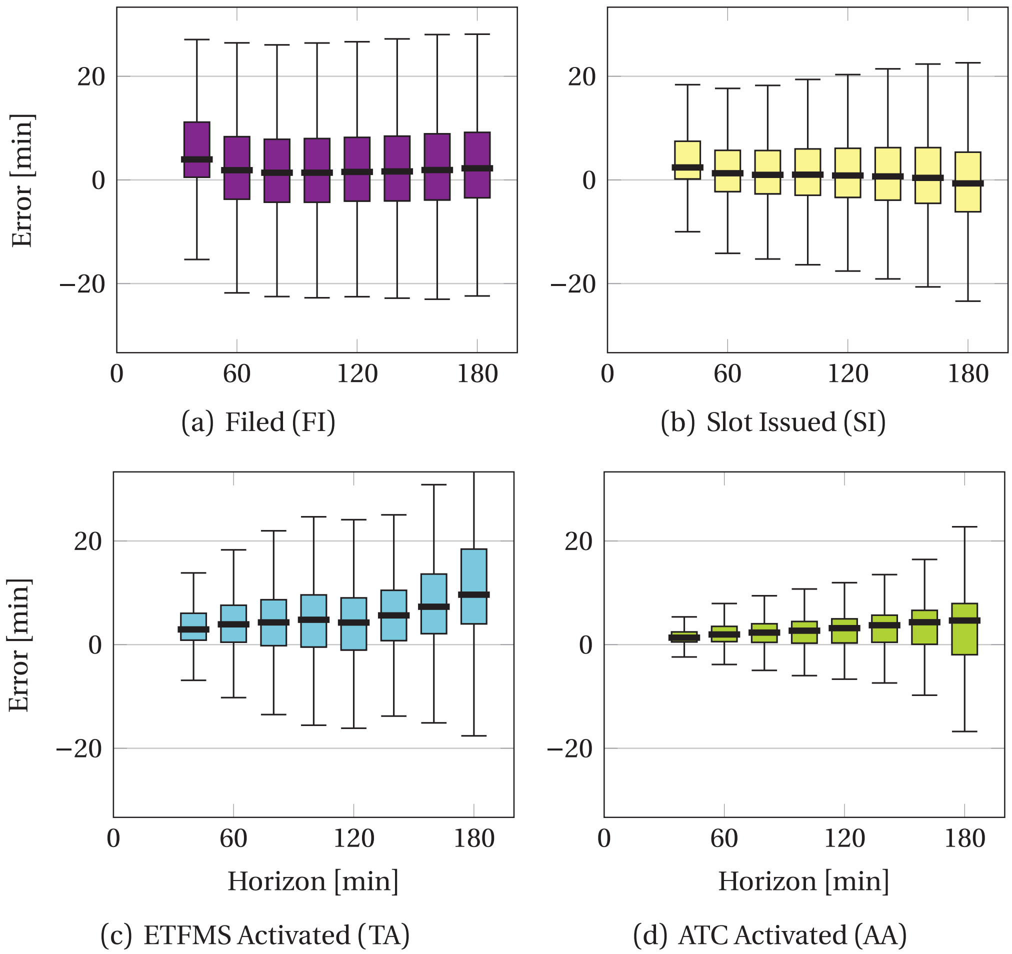


Figure 3.9: Error distribution of predicted time to fly versus remaining time to fly.

Figure 3.9 shows the accuracy of the messages versus the actual remaining time to fly and flight status. Note that data points exist outside the whiskers since the distributions typically have a sharp peak and long tails. The extent of these distributions will be shown further in the chapter.

In general, the prediction medians show a bias toward delay. The bias is smallest when a slot has been issued (SI). The CTOT for these flights forces these flights to depart according to the plan, likely reducing the prediction bias. Most bias toward delay occurs when the NM assumes the aircraft to have departed, but no radar-based report has been received (TA). A likely explanation is that most of those flights have not departed yet and will arrive later. When radar reports come in (AA), the predicted time starts approaching the actual time as aircraft approach the destination, and the uncertainty decreases as expected.

The spread of the prediction errors for the filed status (FI) appears reasonably constant, confirming that deviations mainly depend on events occurring when the aircraft is on the ground. These events do not depend on the time to fly and will not affect the spread of the error with respect to the remaining time. In the other flight statuses, the spread decreases as the time approaches the actual landing time, with airborne reports (AA) showing considerably more precision than other predictions. The latter set also shows stronger convergence of the precision as aircraft approach their destination.

Figure 3.9 plots the error distributions against the actual arrival time, which has been determined afterwards (See Section 3.4.2). In the operation, however, only the predicted arrival time is available. Any uncertainty estimate will need to be determined based on that predicted value. Figure 3.10 provides the same errors as Figure 3.10 but plotted against the predicted time to fly.

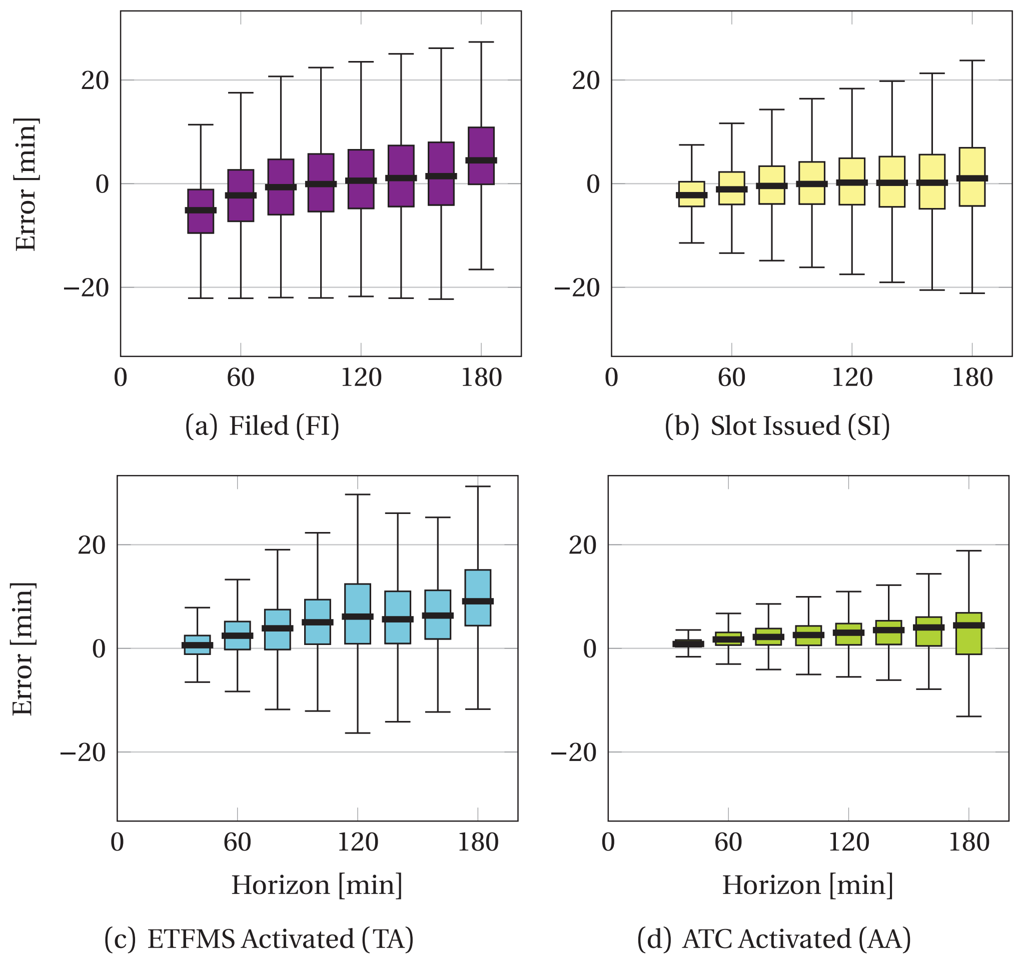
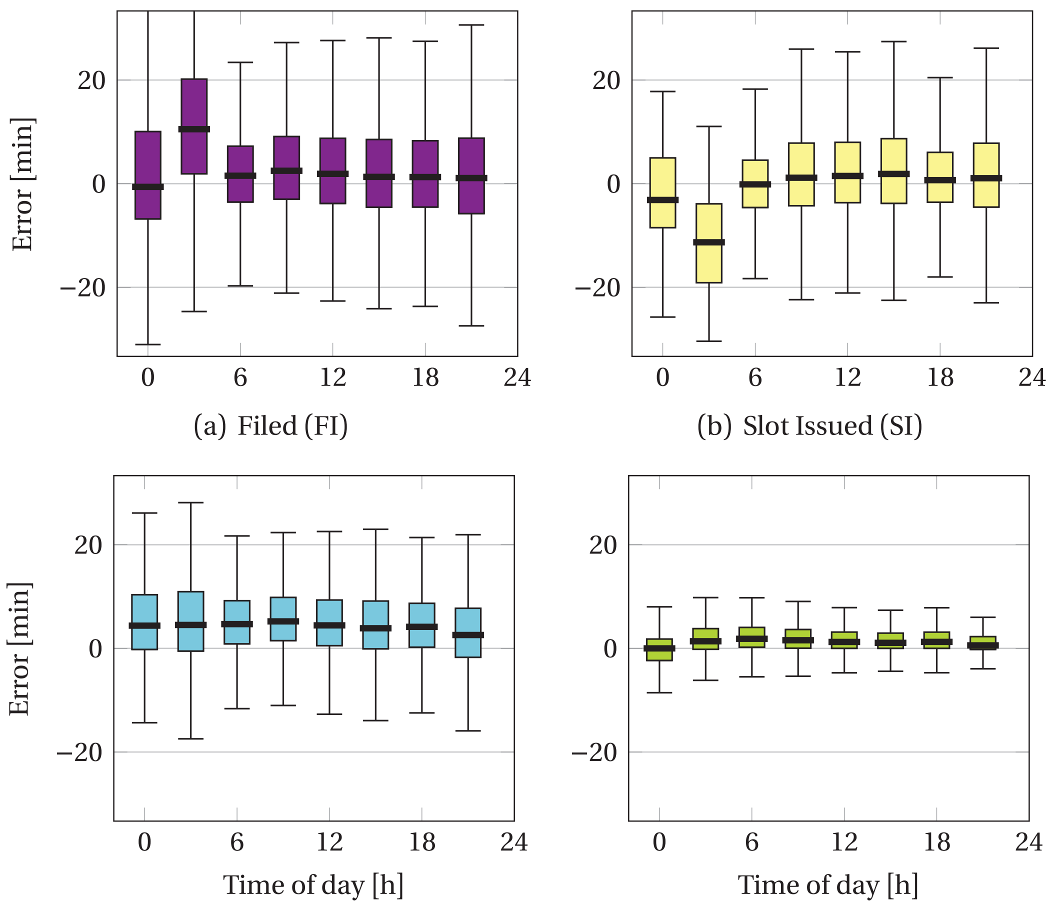


Figure 3.10: Error distribution of predicted time to fly versus predicted time to fly.

In general, the plots in Figure 3.10 reflect similar behaviour as the plots in Figure 3.9. A notable exception is the negative error of the non-airborne prediction

(FI and SI) just before arrival. This is an effect of erroneous predictions being distributed over the prediction horizon (i.e., the error of a flight with a prediction error of 20 minutes at 40 minutes before actual landing is counted at 20 minutes before predicted landing). All graphs show a tendency to predict arrival earlier than actual arrival, which increases with the prediction horizon.

Air traffic during the daytime is considerably heavier than during nighttime. In this period, demand regularly exceeds available capacity in all phases of flight (departure, en-route, and arrival). Limits on the departure cause delays at the departure airport, en-route, and arrival. Demand is balanced by applying restrictions by the NM. These delays affect the accuracy of the expected arrival time, in particular before departure. Figure 3.11 shows this effect, as the spread of predictions for flights that have not been confirmed airborne (FI, SI, and TA) increases for flights predicted to arrive during peak traffic.



(a) Filed (FI)

(b) Slot Issued (SI)

(c) ETFMS Activated (TA)

(d) ATC Activated (AA)

Figure 3.11: Error distribution of predicted time to fly versus scheduled arrival time of day.

3.4.6 Shape of the distributions

Figures 3.10 and 3.11 show the median and the spread of the data. However, due to the large number of data points, the exact shape of the distribution is not always evident. To better understand the effects of different flight parameters on the actual distribution, Figures 3.12 and 3.13 plot the distributions of the errors for subsets of these measurements.

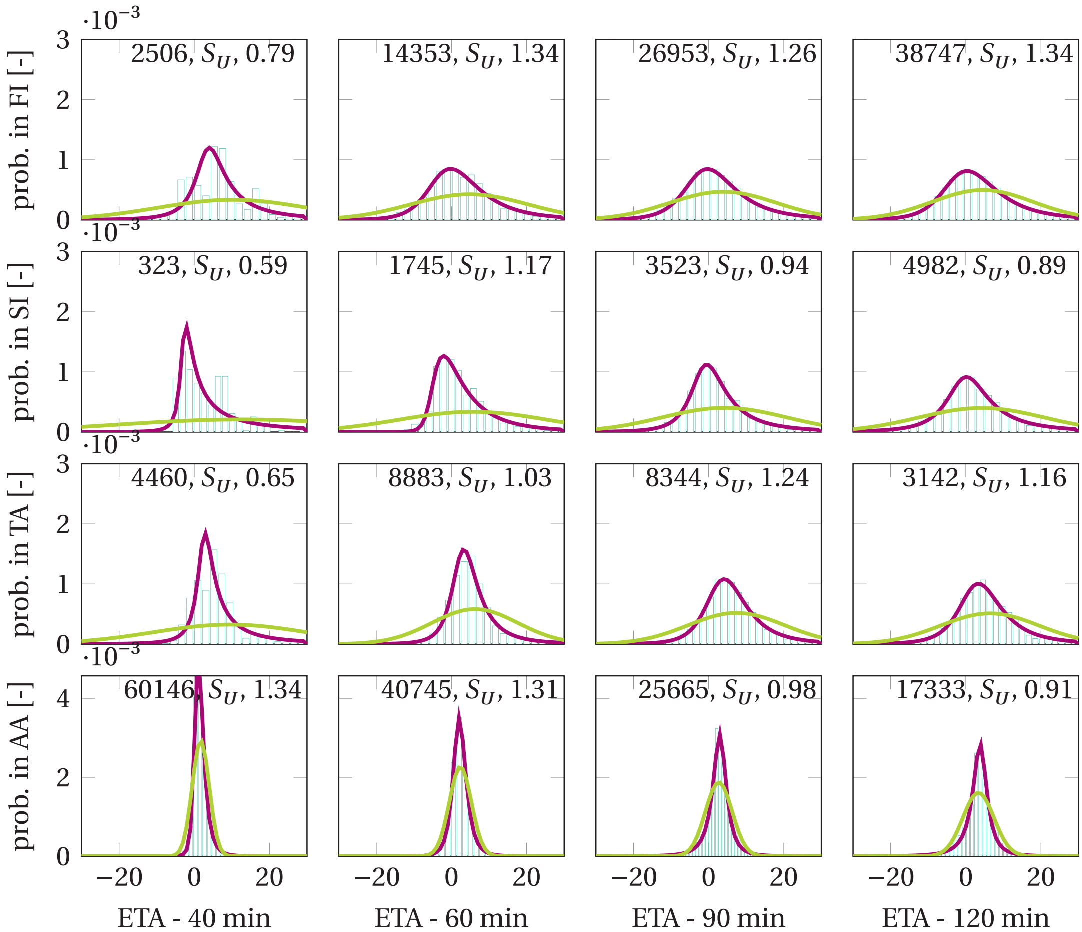


Figure 3.12: Error probability densities (in minutes) against flight status and predicted time to fly for flights scheduled to arrive between 10:30 and 13:30. The numbers in each graph show the number of samples, the type of Johnson distribution and the Kolmogorov-Smirnov test statistic scaled by the number of samples.

The accuracy of predictions, and therefore small spread for airborne flights (AA, Figure 3.10) is confirmed in Figure 3.12: The spread of the data is smaller than 20 minutes. For the other flight statuses, the spread is in the order of one hour.

When not yet airborne (FI, SI), a skew towards delay—left—is visible. The explanation for this error lies in the fact that passenger and cargo considerations make it impractical to depart before the scheduled departure time.

As the prediction horizon increases, the spread of the uncertainty for airborne flights (AA), or for aircraft that are assumed to be airborne (TA), increases as well, as is visible in Figure 3.10. The error of the filed flight plan (FI) does not depend on the prediction horizon, except for very short horizons. Here a limited sample set of likely highly erroneous flights gives a false suggestion of inaccuracy.

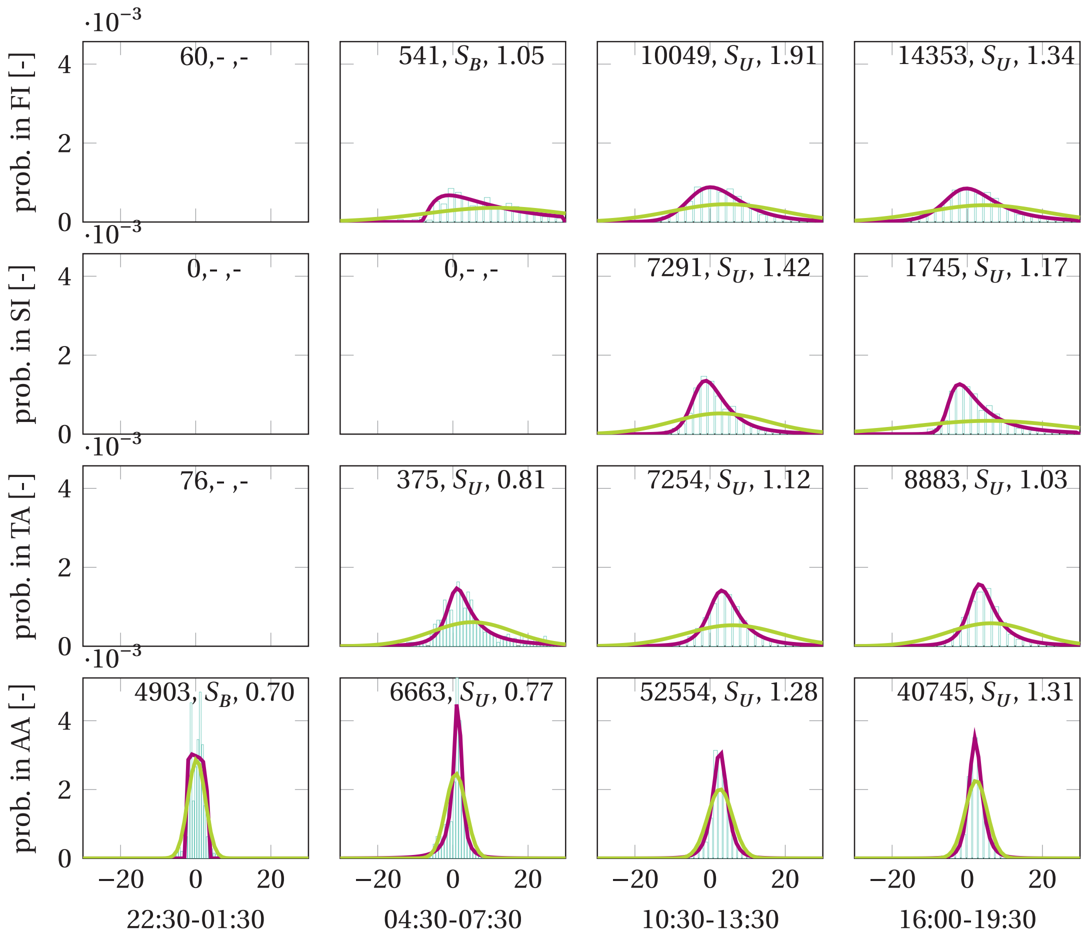


Figure 3.13: Error probability densities (in minutes) against flight status and planned time of arrival at a prediction horizon of 60 minutes. The numbers in each graph show the number of samples, the type of Johnson distribution (S_U or S_B) and the Kolmogorov-Smirnov test statistic scaled by the number of samples. No distribution is calculated when the number of samples is less than 200.

When compared to the scheduled time of arrival in Figure 3.13, the number of flight plans scheduled to arrive in the middle of the night with a flight time of fewer than 60 minutes is small. The lack of data does not allow an analysis of non-airborne flights within this group. The accuracy of the airborne flights at night is high, likely because low traffic levels allow flights to more strictly follow their filed route, and there is no need for ATC to vector them to provide sufficient spacing.

3.5 Modeling distributions

One way to develop a fast modelling technique would be to make the model as simple as possible. For error distributions, such a simple model is a normal distribution. Figures 3.12 and 3.13 show the Probability Density Function (PDF) of normal distribution based on the data in each sample. The graphs show that the normal distribution only provides a reasonable fit for flights confirmed to be airborne (AA, the bottom row).

The variety of different shapes in the distributions makes the versatile Johnson distribution a more suitable candidate for the dataset [22]. This distribution consists of a flexible set of three distributions, which are all transformations of the standard normal distribution. The resulting curve can describe any distribution regardless of mean, standard deviation, skew, or kurtosis.

Five parameters define the Johnson curves. Estimating those parameters for different conditions could support a hypothetical table-based system that provides these parameters as a function of the properties of a received prediction. Subsequently, the distributions could be generated online analytically from the parameters.

3.5.1 Fitting Data

By applying the algorithm developed by Hill et al. [23], a Johnson curve was fit to any dataset with at least 100 samples. The algorithm applies an implementation based on moments [24]. Only the unbounded (S_U) and bounded (S_B) variants were fitted for the available datasets, with the unbounded variant used in most groups.

The bounded variant is applicable when the distribution has no relevant tails. A good example is the case of airborne flights around midnight, 60 minutes before arrival, as shown in Figure 3.13. The number of errors larger than 5 minutes is so small that they are negligible. The third variant—the log-normal variant, S_L—forms the boundary case between the two other variants [22]. Since this is an edge case, its occurrence in a random dataset is unlikely.

The fitted curves are shown in Figures 3.12 and 3.13. The graphs show that the fitted distributions are close to the histograms of the dataset.

The goodness-of-fit was validated using the Kolmogorov-Smirnov test statistic. This statistic is the maximum absolute difference between the Cumulative Distribution Functions (CDFs). The Kolmogorov-Smirnov test assumes, however, that the two compared CDFs are independent. The distribution has been determined by fitting to the same data. While the test value serves as an indicator of goodness-of-fit, it cannot be compared to a critical value to test significance.

A bootstrap resample process enables the bounding of the significance level of the difference between the CDFs. From each subset, a random sample of 100 points is selected with replacement. From this sample, the Kolmogorov-Smirnov statistic to the CDF of the Johnson distribution is calculated. The process is repeated 10,000 times to determine the confidence interval for the statistic. Typically, the 95% confidence interval of the statistic reaches a maximum of 1.65, with 1.95 being the cut-off value to statistically reject the hypothesis that the distributions match at a 0.001 confidence level.

The statistical tests show that the fitted function is often statistically different from the dataset. For application purposes, however, the difference in the CDF—which is not multiplied by the square root of the number of sample points—will be in the order of 1% making decisions based on the CDF likely to be accurate.

3.6 Validation

Since the Johnson distributions have algebraic equations for both the PDF and the CDF, a look-up table with the parameters provides a fast way of calculating these curves as needed. This supports real-time application as needed in arrival management. The method's suitability then depends on the robustness of the fitting technique and the effects of the discrete bins on the continuously changing distributions.

To test the robustness of this technique, a process akin to the bootstrap process was applied to test the goodness-of-fit of an estimated Johnson curve. A subsample of 100 points from each bin was selected randomly and used to fit a Johnson-curve. Subsequently, goodness-of-fit was determined for 100 different points from the same bin. The number of selected points was based on the number of points in the typical bin to allow testing of most bins.

This process was repeated 50 times to determine the confidence bounds for each subset of data. The maximum value of the confidence interval of the mean for Kolmogorov-Smirnov statistic varies up to 2.02 for AA flight statuses and less than 1.95 for the other statuses, at which point the two distributions cannot be proven to be different ($p<0.001$). This shows that, while sampling affects validity, the effect on the CDF is unlikely to be substantial.

If an online display would show uncertainty, gradual changes are preferable to discrete changes. While the change of a flight status will always present a discrete change, the change of prediction horizon is a continuous process. Due to the use of binned groups, a discrete change will occur at the transition of one horizon bin into the next. The selection of bin width for such continuous parameters is a compromise between sufficient data points to generate an accurate distribution and the magnitude of the step change of the distribution shape at the bin transition.

In Chapters 4 and 5 we will propose a display that shows uncertainty. The main visual feature of uncertainty is the width of a flight's arrival time PDF [25]. Analysis of the width of the 95% confidence interval over each transition shows the effect of the transitions. Table 3.2 provides the width of the PDF for airborne flights (AA) different times of the day.

At an interval of ten minutes, the step size can be up to half of its width (e.g., from 50 to 40 minutes in the second row). In a visual context, these would cause considerable visible changes for an otherwise continuous process.

Table 3.2: Effect of horizon bin transition on the width of the PDF for flights with flight status AA.

STA	Width of central 95th percentile of the PDF at different horizons [min].						
	90	80	70	60	50	40	30
22:30 - 01:30	8	7	6	6	5	6	3
01:30 - 04:30	15	14	15	14	11	7	3
04:30 - 07:30	18	16	15	12	10	8	4
07:30 - 10:30	15	15	13	11	9	7	4
10:30 - 13:30	15	15	14	11	9	7	3
13:30 - 16:30	16	16	16	12	10	7	4
16:30 - 19:30	17	16	15	11	10	7	4
19:30 - 22:30	13	13	13	12	10	8	4

3.7 Discussion

This chapter demonstrates a method for predicting arrival time uncertainty using currently available estimates. The use of FUMs implies, however, that the proposed method inherits the prediction errors included in those messages. Two of the most important error sources are the assumed route in the destination airspace and the effective resolution of the ETA in those messages.

Since the reported landing times were used as the reference value, any difference between the assumed arrival route and the actual route will result in a prediction error. In predicting, the NM assumes that the aircraft follow published approach routes. At Schiphol Airport, aircraft follow these published routes only during nighttime. During day time, ATCOs will direct the traffic. The resulting uncertainty in landing time is not meaningful in the context of AMAN, as the controllers will use AMAN as an input in directing the traffic, amongst other information. By reducing the error at 20 minutes before arrival, this error was negated

(Figure 3.8). However, the resulting errors are not fully representative of the actual prediction error at that time. Future concepts may address this inaccuracy by comparing times over particular waypoints.

Since NM provides an update only when the actual progress of the aircraft deviates more than five minutes from the prediction, errors below this limit are not taken into account. They are, therefore, not present in the data. This problem is partially addressed by the interpolation, thus generating sampled errors of the accurate predictions between the first prediction and the arrival.

The interpolation method does introduce a new source of potential errors. The assumption of a linearly changing error between two reports can theoretically introduce some linearity in the modelled uncertainty. On the other hand, the variety in the number of error reports for a single flight and varying errors themselves counteract the linearity.

3.7.1 Other airspaces

Since the data demonstrate that departure uncertainty is the primary cause of uncertainty in the arrival time, both the airline departure accuracy and the delays at airports generate the majority of the uncertainty. These two factors account for 20-40% and 5-10% of arrival time deviation respectively and depend on the airline, and the origin airport [18]. This effect is exacerbated through reactionary delay, which is the delay caused by the delay of the inbound aircraft or inbound transfer passengers. These will cause airlines with numerous rotations between two airports or large numbers of connecting passengers to have extra arrival time uncertainty. Therefore, the arrival time accuracies of flights strongly depend on the city pairs and, thus, the arrival airports. A model estimated for one airport will not necessarily apply to other airports and will require an estimate for each particular case.

3.7.2 Other data sources

The FUM was selected since it was available and recorded at LVNL. Only the four most common flight statuses in those FUMs were used. The NM nowadays—in 2022—provides a more detailed report in the form of the EFD message. This message uses the same data source and is provided under the same conditions. However, the message provides more detailed routing information and, therefore, more possible reference data to determine the prediction error. Future research should evaluate the additional information to improve the accuracy of the uncertainty prediction.

A-CDM provides more details on the flight's progress towards take-off. Since the data show high uncertainty mainly on departure, the inclusion of A-CDM information might allow receiving more detailed flight states. Unfortunately, the number of origin airports that provide A-CDM information was too low to provide sufficient data for this research. Chapter 6 will perform an initial analysis on the effects of using the knowledge that the departure airport provides such information.

3.7.3 Bin selection

The parameters of interest and their respective bin widths in this study are based on the number of available data points per bin. The current selection of properties is based on an initial analysis, combined with the likelihood that those properties would have an influence. Section 3.4 suggests a further set of parameters that may be selected. The potential list of properties provides further options for investigation. For example, it may be worthwhile to distinguish between departure airports by separating flights from the most prevalent airports and grouping the remainder. This approach would enable modelling the uncertainties due to specifics at those departure airports. Further detailing can come from information available in other predictions, such as A-CDM messages or EFD.

For simplicity, horizon bin sizes were selected at regular intervals. Section 3.6 describes these bin widths and shows that the transition from one bin to the next one can cause a considerable change in the width of the PDF. Especially in visual presentation, such instability may be an issue in acceptance by ATCOs. The size of bins may well be adjusted based on the other criteria to reduce this behaviour. For example, uncertainty as a function of Scheduled Time of Arrivals (STAs) is likely to be more variable during the daytime than at night, particularly for airports with wave-like traffic densities such as Amsterdam.

3.8 Conclusion

This chapter presents a method for predicting uncertainty in arrival time at longer horizons using readily available data at LVNL. By fitting Johnson curves to errors derived from FUMs generates that can more precisely describe the shape of the distribution, in particular skew and kurtosis of the uncertainty. By using the Johnson distribution from tabulated data, based on several parameters of the received prediction, an error distribution can be calculated fast and is therefore suitable for on-line applications such as displays for ATCOs.

3.9 Bibliography

[1] N. Hasevoets and P. Conroy, “Arrival Manager - implementation guidelines and lessons learned,” EUROCONTROL, Brussels, Belgium, Tech. Rep., 2010.

[2] A. Barff, B. Favennec, P. Conroy, *et al.*, "SESAR P05.06.04 - D28 - Preliminary OSED ed. 00.01.01," SESAR Consortium, Tech. Rep., 2012.

[3] J. Bronsvoort, G. McDonald, M. Paglione, C. Garcia-Avello, I. Bayraktutar, and C. M. Young, "Impact of missing longitudinal aircraft intent on descent trajectory prediction," in *30th IEEE/AIAA Digital Avionics Systems Conference*, Seattle, WA, USA, Oct. 2011.

[4] "SESAR concept of operations at a glance ed 02.00.00," SESAR JU, Tech. Rep., 2011.

[5] JPDO, *NextGen integrated work plan: A functional outline*, 2008. [Online]. Available: `http://www.jpdo.gov` (visited on 03/01/2013).

[6] J. Bronsvoort, G. McDonald, R. Porteous, and E. Gutt, "Study of aircraft derived temporal prediction accuracy using FANS," in *Proceedings of the 13th Air Transport Research Society (ATRS) World Conference 2009*, Abu Dhabi, UAE, 2009.

[7] E. R. Mueller and G. B. Chatterji, "Analysis of aircraft arrival and departure delay characteristics," in *AIAA's Aircraft Technology, Integration, and Operations (ATIO) 2002 Technical Forum*, Los Angeles, CA, USA, Oct. 2002.

[8] M. Whiteley, "PHARE advanced tools tactical load smoother final report," EUROCONTROL, Brussels, Belgium, Tech. Rep. November, 1999.

[9] D. Nicholls, P. Marti, P. Battino, S. Pozzi, and V. Barsotti, "CARE innovative action presenting uncertainty to controllers & pilots," EUROCONTROL, Brussels, Belgium, Tech. Rep. December, 2004, pp. 1–53.

[10] T. Mueller, J. Sorensen, and G. Couluris, "Strategic aircraft trajectory prediction uncertainty and statistical sector traffic load modeling," in *AIAA Guidance, Navigation, and Control Conference and Exhibit*, Reston, VA, USA, Aug. 2002.

[11] G. Tobaruela, P. Fransen, W. Schuster, W. Y. Ochieng, and A. Majumdar, "Air traffic predictability framework - Development, performance evaluation and application," *Journal of Air Transport Management*, vol. 39, pp. 48–58, 2014.

[12] R. A. Coppenbarger, "Climb trajectory prediction enhancement using airline flight-planning information," in *AIAA Guidance, Navigation, and Control Conference*, San Diego, CA, USA, 1999.

[13] G. Sölveling, "Stochastic programming methods for scheduling of airport runway operations under uncertainty," Dissertation, Georgia Institute of Technology, Jun. 2012.

[14] "Traffic Review 2013," Amsterdam Airport Schiphol, Amsterdam, The Netherlands, Tech. Rep., 2013.

[15] C. Wanke, S. Mulgund, D. Greenbaum, and L. Song, "Modeling traffic prediction uncertainty for traffic management decision support," in *AIAA Guidance, Navigation, and Control Conference and Exhibit*, Reston, VA, USA: AIAA, Aug. 2004.

[16] H. Koolen and S. Smidt, "Flight Progress Messages," EUROCONTROL Network Manager, Brussels, Belgium, Tech. Rep., 2014.

[17] T. Morisset and A. Odoni, "Capacity, delay, and schedule reliability at major airports in Europe and the United States," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2214, pp. 85–93, Dec. 2011.

[18] "Delays to air transport in Europe - annual 2013," EUROCONTROL CODA, Brussels, Belgium, Tech. Rep., 2014.

[19] "All-causes delay and cancellations to air transport in Europe - 2014," EUROCONTROL CODA, Brussels, Belgium, Tech. Rep., 2015.

[20] "Performance review report 2013," Performance Review Commission, EUROCONTROL, Brussels, Belgium, Tech. Rep. May, 2014.

[21] "Performance review report 2014," Performance Review Commission, EUROCONTROL, Brussels, Belgium, Tech. Rep. May, 2015.

[22] N. L. Johnson, "Systems of frequency curves generated by methods of translation," *Biometrika*, vol. 36, no. 1/2, pp. 149–176, Jun. 1949.

[23] I. D. Hill, R. Hill, and R. L. Holder, "Algorithm AS 99: Fitting Johnson curves by moments," *Applied Statistics*, vol. 25, no. 2, pp. 180–189, 1976.

[24] D. Jones, "The Johnson curve toolbox for Matlab: analysis of non-normal data using the Johnson system of distributions," College of Marine Science, University of South Florida, St. Petersburg, FL, USA, Tech. Rep., 2014. [Online]. Available: `http://www.marine.usf.edu/user/djones/`.

[25] M. Tielrooij, C. Borst, D. Nieuwenhuisen, and M. Mulder, "Supporting arrival management decisions by visualising uncertainty," in *Proceedings of the SESAR Innovation Days*, D. Schaefer, Ed., Stockholm, Sweden, 2013.

Chapter 4

Visualizing uncertainty in arrival management

Chapter 2 explains that uncertainty limits the validity of arrival management decisions and proposes applying knowledge on uncertainty in the decision-making process. Chapter 3 provides a means to predict the uncertainty associated with the arrival time of a particular flight. Irrespective of whether the decisions are made by a human operator or an automated process, a Human-Machine Interface will be needed to execute or monitor the decisions, respectively. This chapter proposes a visual interface that provides information on the accuracy of an Estimated Time of Arrival. The basis of this chapter is an existing publication. This may cause some repetition of previous chapters.
This and the following chapter were developed in parallel with the work in Chapter 3. The experiments in these two chapters use a simplified uncertainty model that has no relation to the model proposed there. This allowed for a more controlled experimental setting and limited dependency on progress in both streams.

This chapter is based on the following publication:

Paper Title	Supporting arrival management decisions by visualising uncertainty
Authors	M. Tielrooij, C. Borst, D. Nieuwenhuisen and, M. Mulder
Published in	Proceedings of the SESAR Innovation Days, November, 2014, Stockholm, Sweden

.

4.1 Introduction

Many Air Navigation Service Providers (ANSPs) nowadays use Arrival Managers (AMANs) to plan the arrival times of inbound aircraft and thereby balance demand to the available landing capacity. These systems provide support to the sequence manager (US: Traffic Management Coordinator) in deciding how to modify the 4D trajectory of inbound aircraft when these are predicted to arrive with too little spacing between them.

When aircraft are assumed to fly the optimal trajectory from an Airspace User's (AU) perspective, deviations from these trajectories should be kept to a minimum. Any remaining change in trajectories should be performed as early as possible for optimal flight efficiency. For example, a smaller speed increase over a longer flight time is much more fuel-efficient than a more substantial speed increase over a shorter time while ultimately achieving the same difference in time [2].

Currently, the horizons at which AMAN systems monitor and influence traffic are typically limited to 20 to 30 minutes (or 150-200 NM) by three factors [3]:

1. The availability of information on the predicted arrival time of aircraft, for example, due to the limit of radar surveillance,
2. the ability to influence the aircraft, for example, because of the boundaries of Flight Information Regions (FIRs), and
3. the reliability of the predicted arrival times.

Future operational concepts, such as those proposed in Single European Sky ATM Research (SESAR) and its counterpart in the United States: Next Generation Air Traffic Management System (NextGen), foresee an increase in the planning horizon to increase the efficiency and predictability of operations. Depending on the concept, the future horizons are expected to be at 200-500 NM [2], [4], [5]. The three limitations on the AMAN horizon need to be overcome to achieve this increase.

Current developments address the first two constraints: System Wide Information Management (SWIM) should enable continuous sharing of all relevant information concerning a flight between all involved actors [6], [7]. And, through SWIM, different ANSPs can share requirements on a trajectory (such as an arrival time planned by AMAN) and execute such shared requirements. The third constraint *prediction uncertainty* is expected to reduce but is unlikely to disappear altogether. Increasing the AMAN horizon will then require other ways to perform arrival planning in the presence of uncertainty.

This chapter proposes a display that provides operators with information on the uncertainty in the arrival time. By doing so, it is envisioned that the controller can actively decide whether to adjust a trajectory or to delay a decision until the arrival time is more accurate.

Section 4.2 and 4.2, respectively, discuss the approach and reviews the abstraction hierarchy as described in Chapter 2. Section 4.4 then considers the effects of uncertainty on that abstraction hierarchy. Using that framework, Section 4.5 adds elements to the AMAN-time line to show uncertainty and its relevance to the schedule. Sections 4.6 through 4.9 describe the setup and results of two human-in-the-loop experiments in which novice users planned an arrival schedule under uncertainty. The last two sections discuss these experiments and their conclusions.

4.2 Approach

Most of the current AMAN interfaces use a vertically moving time line showing the expected or planned time of arrival of all approaching aircraft [8]. An example of such a time line is presented in Figure 4.1. This presentation reduces the 4D trajectory information to a one-dimensional spacing problem as it exists to the sequence manager. The sequence manager is not separating aircraft in 3 dimensions as radar controllers do, but rather plans the spacing between consecutive aircraft. However, the single-dimensional presentation hides the complexity of the total system under control. It does not clarify how decisions on individual aircraft affect the achievement of other goals of the system. For example, determining the required delay for the last aircraft in a group requires sequential spacing of *all* involved aircraft and then determining how far the last aircraft has deviated.

The task of planning arrival times has several factors which make it a complex task:

- Many of actors (aircraft, Air Traffic Controllers (ATCOs), airport operators),
- time delay,
- potential interactions between aircraft in a sequence, and
- uncertainty in the provided information.

The task often requires expert knowledge and results in a high task demand in integrating all factors. In most cases, sequence managers are experienced ATCOs. These controllers develop an arrival plan using their knowledge and experience and the information on predicted arrival time. In these decisions, they will use not only the Estimated Times of Arrival (ETAs) but also information such as runway availability, weather, team competency, and airspace capacity.

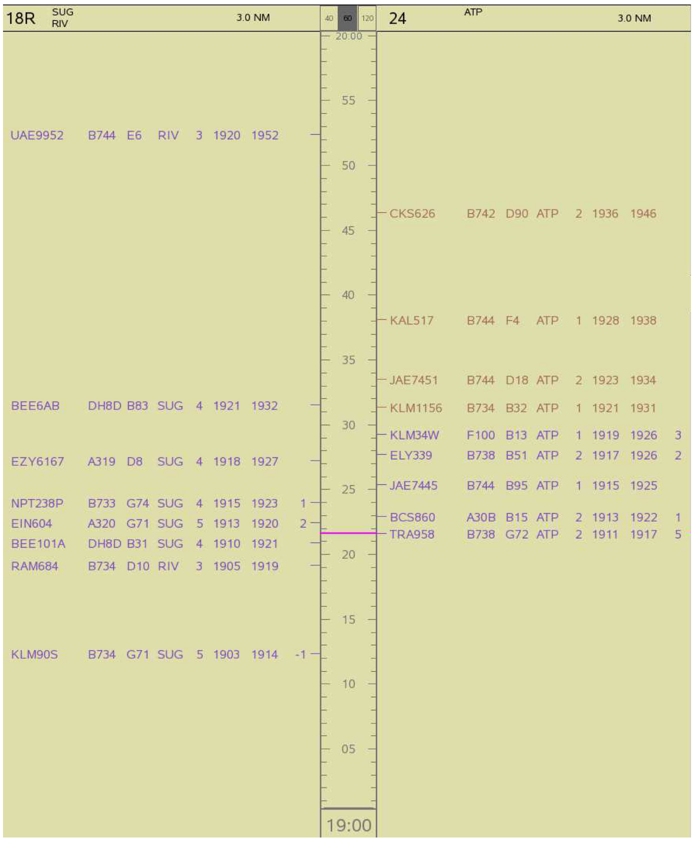


Figure 4.1: Typical AMAN time line as used by LVNL (inverted colours). Source: LVNL

The uncertainty in the predictions is not available to the operators. Operators may have experienced at what horizon and in what situations the AMAN HMI no longer provides support and hence apply a—limited—model of uncertainty in their mental model of the operation. To incorporate uncertainty to the many other factors in the arrival management process, this chapter will use a structured approach to determine how to visualise the arrival management environment and the uncertainty:

1. To discover how uncertainty affects the operator's decision making and the effect of uncertainty in the Decision Support Tool (DST) on the operation, the work domain (i.e., the AMAN process) is analysed using Rasmussen's Abstraction Hierarchy [9]. This analysis uses the abstraction hierarchy introduced in Section 2.7.2. In this section, the description will be limited to the relevant components of present-day operational AMAN interfaces.

2. Once the abstraction hierarchy is established, the uncertainty of the properties of each element is determined.

3. Using the functional relationships, the effects of the uncertainty can then be

related to the objectives of the system. This step explains how the various origins of uncertainty propagate through the AMAN system.

4. Finally, the commonly used time line is extended stepwise by using the knowledge gained in defining the abstraction hierarchy.

4.3 Work domain analysis

Figure 2.11 in Chapter 2 provides an analysis of the work domain using the abstraction hierarchy. This description uses the SESAR Key Performance Indicators (KPIs) as the functional purposes of the AMAN system [3], [10]. In this section, the abstraction hierarchy is modified to focus on present-day operations and limit itself in the scope of the work domain. The resulting hierarchy is shown in Figure 4.2.

In this figure, the colours of each element indicate the relative uncertainty of the information in that element. The uncertainty of each of the elements will be explored in Section 4.4. For clarity, the names of elements in the hierarchy are emphasised in the text.

Note that the original *Functional Purposes* in Figure 2.11 included all SESAR KPIs to which AMAN can possibly contribute. The following analysis focuses on the one-dimensional planning aspect, in which the constraint on traffic is the available capacity at an airport. The analysis does not include the spatial relation between aircraft and their respective routes to the airport. As a consequence, the effects on the workload of upstream controllers are disregarded. The same holds for the effects of AMAN decision making on *Cost-Effectiveness* as it derives from the workload.

The environmental impact of a planning decision is determined by the selected speed of an aircraft (i.e., the amount of fuel spent per distance) and the amount of time that an aircraft is in the air (i.e., the amount of time the engines are burning fuel). Therefore, this analysis assumes that the AU's objectives—represented in *Efficiency*—are coupled with the *Environment* KPI.

Finally, predictability is not part of the AMAN decision-making process but rather a consequence of the use of AMAN. Therefore, the analysis will disregard *Predictability*.

Uncertainty is introduced as descriptions of the *Physical Form*. As those properties become more uncertain, the resulting outcomes are also less certain. The uncertainty in the physical form propagates upward through the abstraction levels. High uncertainty in the *means* result in higher uncertainty in the *ends*. At the AMAN horizon under study, the relative size of prediction errors are largest for the ETA (See Chapter 3). Through the means-end chain, the high uncertainty in the ETA leads to high uncertainty for all elements at the abstract function level.

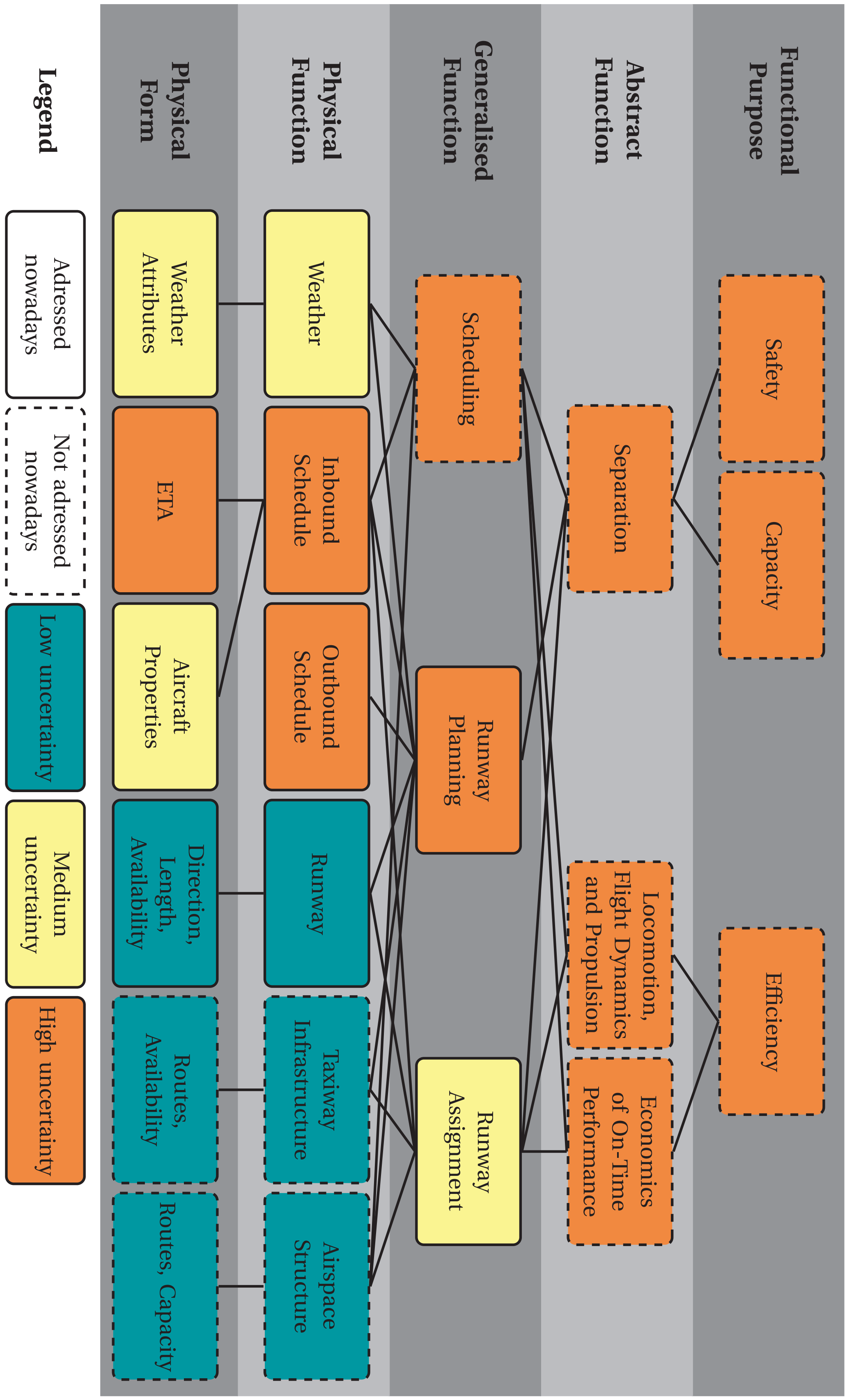


Figure 4.2: Reduced abstraction hierarchy of the AMAN work domain indicating the relative uncertainty originating as described in Chapter 4.4.

4.4 Uncertainty

Now that the abstraction hierarchy is established, the origins and magnitude of uncertainty in the process can be analysed further and superimposed on the abstraction hierarchy description. Chapter 3 discussed several sources of uncertainty in the predicted ETA. Furthermore, this section will consider uncertainties originating from other elements of the work domain. The initial distinction is made by comparing the relative magnitudes of the uncertainties. Figure 4.2 shows the relative uncertainty using the colour of the elements on the hierarchy.

4.4.1 Physical form

The uncertainty in ETA propagates into an uncertainty in the resulting *Inbound Schedule*, as the arrival times underlying the schedule are uncertain. As a consequence of the uncertain ETAs, the optimal schedule from a capacity and efficiency perspective also becomes uncertain. At a two-hour horizon, the uncertainty can vary between 10 and 40 minutes (as found in Section 3.4).

Uncertainty in the ETA includes potential errors in wind speed and direction. However, weather conditions also influence the runway capacity, both through wind and visibility. As shown in Table 2.2, the headwind determines the number of aircraft per hour given a specific spacing requirement. Furthermore, larger spacing intervals are applied when visibility is low. While some error exists in weather predictions, the variation in time, and therefore uncertainty in the horizon under consideration, is relatively small compared to the uncertainty in ETA.

Uncertainty in *Aircraft Properties* influences the schedule directly in choices on the flight's schedule. As with the weather, however, this is directly covered in ETA uncertainty. Choices in aircraft operation during landing and on the runway also influence capacity. These choices include the landing speed—based on the actual aircraft weight—and the selected procedures on the runway which determine how quickly the aircraft vacates the runway. Chapter 3 demonstrated that the accuracy of ETAs is in the order of several minutes. The magnitude of the variation in runway occupancy and landing speed is in the order of tens of seconds. The effect of these uncertainties is relatively small when compared to the uncertainty in ETA.

The availability and capacity of the airport infrastructure vary. However, unplanned closures are uncommon; therefore, the uncertainty is very low compared to all other aspects.

4.4.2 Physical function

Section 3.4 showed considerable uncertainty for aircraft that did not yet leave the departure airport. Therefore, the uncertainty of the departure schedule is comparable in magnitude to the uncertainty of the arrival schedule. The errors for the departure schedule realise slightly later. As demonstrated, the increased prediction horizon further increases uncertainty.

4.4.3 Generalized function

At most airports, the structure of the airspace determines runway assignment. Currently, at Amsterdam's Schiphol airport for example, three Initial Approach Fixes (IAFs) feed two runways during an inbound peak. Two of those IAFs are permanently assigned to their set runway. Aircraft from the third fix generally use one runway and are only scheduled occasionally to the other runway. *Runway Assignment* is therefore primarily dependent on the airport infrastructure and only slightly affected by uncertainties in the inbound and outbound schedule.

Planning when to open which runway for arrivals or departures strongly depends on the respective schedules. Since both schedules are highly uncertain, the runway planning process also becomes uncertain. Therefore, at Amsterdam, the decision on when to switch runway configuration can only be made about 20 minutes before the first aircraft lands or departs from the newly opened runway.

4.4.4 Purpose of AMAN

The uncertainty in *Scheduling* propagates to all elements of the *Generalized Function* level. This propagation implies that uncertainty in the *ETA* ultimately affects the ability to achieve all of the purposes of the AMAN process. However, describing the effect of uncertainty at the *Functional Purpose* level is more complex.

When scheduling, uncertainty does not result in lower or higher quality schedules. Instead, the presence of uncertainty makes scheduling harder. Delaying the scheduling process lowers the number of required corrections. However, the later decision-making may reduce flight efficiency as described in Section 4.1.

To improve the ability to achieve the purposes of AMAN, the uncertainty on ETA needs to be provided to the operator. However, since the effect of uncertainty is unclear, its relation to the higher abstraction levels also needs to be visualised.

4.5 Visualisation

Looking at the top three strata of the abstraction hierarchy in Figure 4.2, the objective of the new AMAN display will be twofold: To provide information on the uncertainty associated with the ETA and its effect on the purposes of the system, and second, to provide support in planning even while uncertainty is too high to warrant assigning definitive arrival times.

The time line presentation in Figure 4.3(a) is a very suitable way of presenting arrival times. It is an analogue display that enables immediate visual inspection of spacing between aircraft while simultaneously allowing for meaningful direct manipulation of the arrival schedule (i.e., assigning arrival time by moving aircraft on the line). Therefore, the time line forms the base diagram.

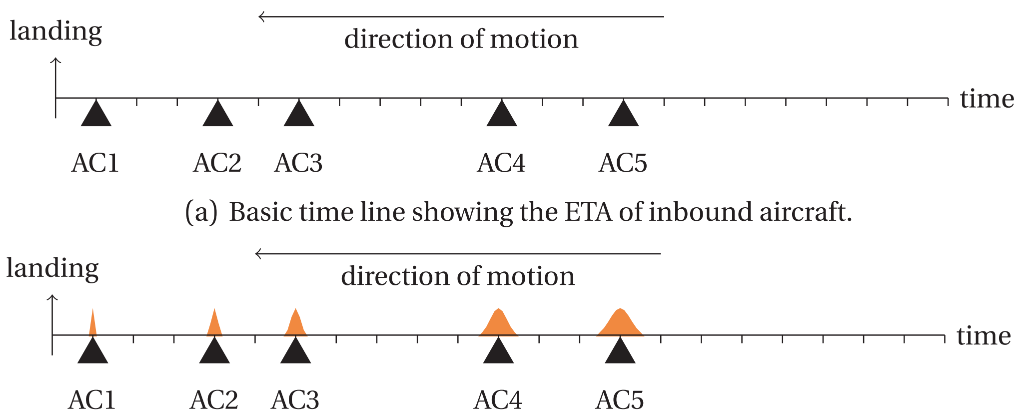
In the following subsections, the abstraction hierarchy in Figure 4.2 guides the development of a new display, starting with the traditional time line. Each addition is clarified by highlighting the area of the abstraction hierarchy that is addressed.

4.5.1 Basic time line: ETA and schedule

On the hierarchy, the time line directly provides the *ETA* and the resulting *Inbound Schedule*. This is indicated on Figure 4.4 as a dashed area.

Section 2.5 describes that very few current time line diagrams present the spacing requirement. So, while the diagram displays the inbound schedule, it does not provide further support in scheduling. As such, it does not directly support the purposes of the system.

Figure 4.3(b) shows that uncertainty in arrival time can be visualised directly on a time line as the Probability Density Function (PDF) of the arrival time. This approach would directly provide information on the uncertainty to the operator. However, it would require the operator to interpret the information and determine the effects of the uncertainty on achieving the objectives.



(a) Basic time line showing the ETA of inbound aircraft.

(b) Extending the basic time line with the uncertainty on ETA.

Figure 4.3: Basic time line showing ETA and uncertainty.

4.5.2 Separation and capacity

To develop support for interpretation of the effects of uncertainty on the purpose of the system, we take a step back and evaluate the use of the time line without any uncertainty using the abstraction hierarchy: At the *abstract function* level, the key missing element in the display is the required spacing between two flights.

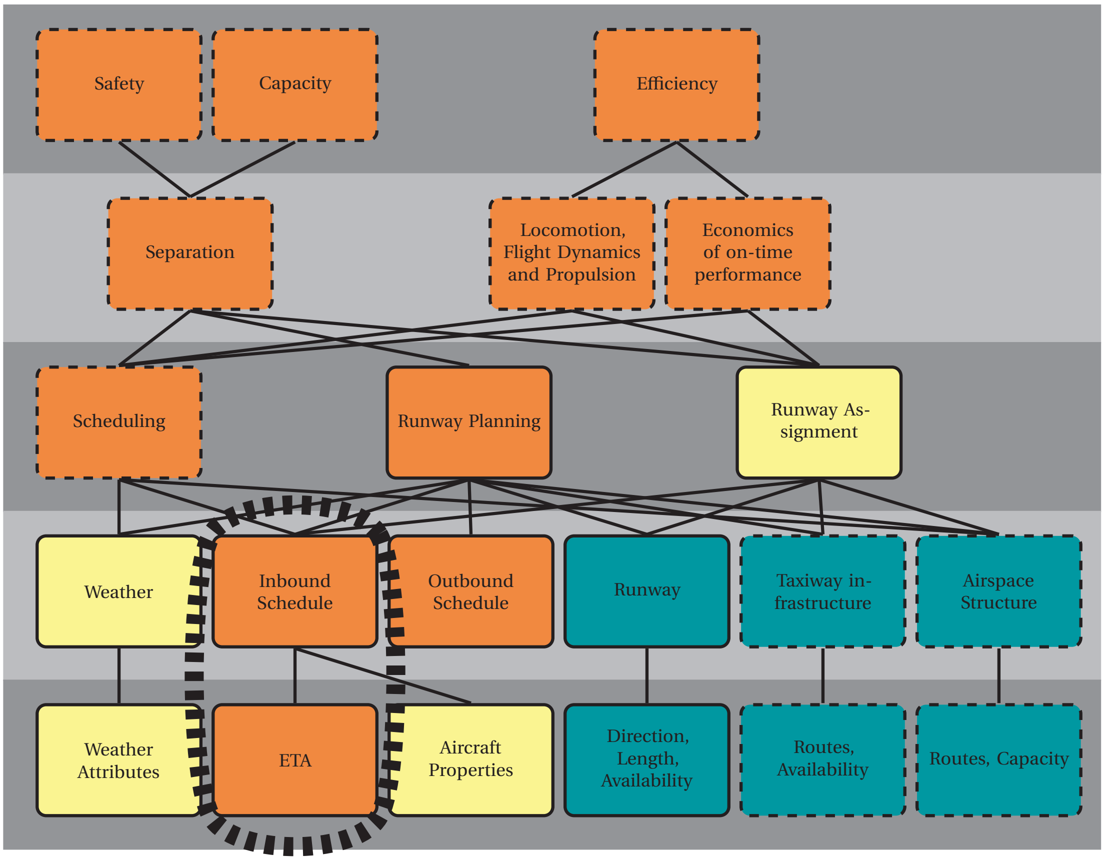
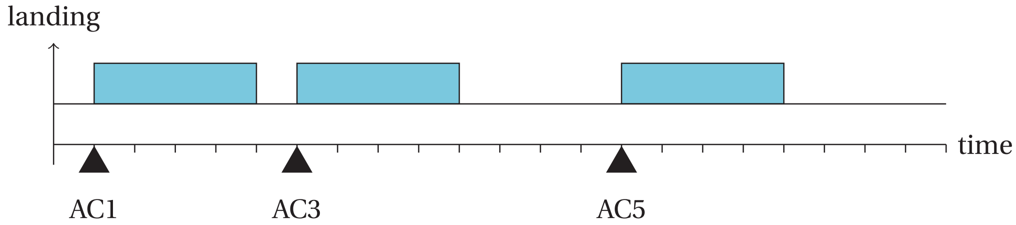


Figure 4.4: Visualisation of ETA on the abstraction hierarchy.

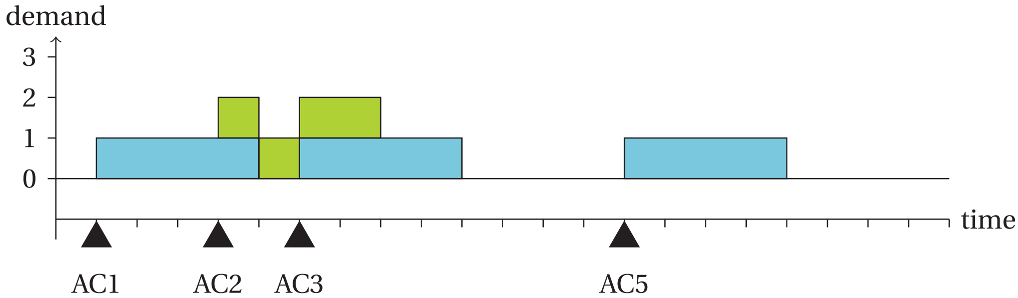
When spacing is defined in time, a bar along the time line can directly visualise the required spacing. This approach is used in several experimental time lines (See Chapter 2). Figure 4.5(a) shows the required spacing for three aircraft. This presentation visualises the relation between ETA and the *safety* objective (See Figure 4.6).

When spacing intervals are expressed in time, the interval represents the number of seconds an aircraft occupies a limited resource (i.e., the runway). When the blocks in Figure 4.5(a) are defined one “unit” high, the areas represent the demand that the aircraft imposes on the runway. This property allows superimposition of the areas of multiple aircraft to present the total demand on the resource. Figure 4.5(b) shows the ‘stacked’ demand.

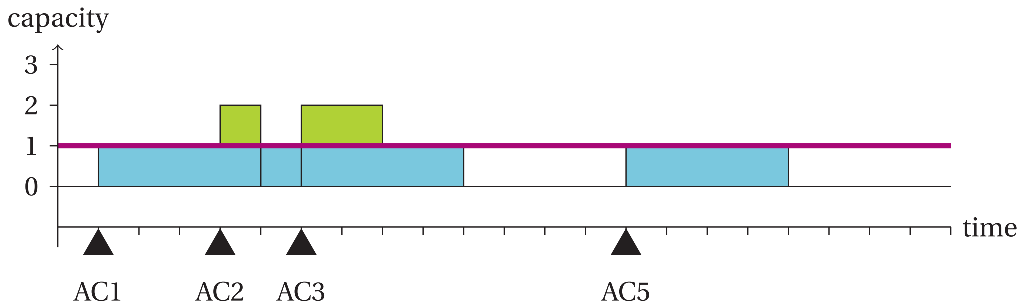
Finally, the total demand for the resource (i.e. the number of flihgts) needs to be compared to the capacity that the resource (the number of runways) currently provides. In the case of a single runway, the capacity offered by that runway is one unit each second. On the time line, the capacity is a line at the number of runways, shown in Figure 4.5(c).



(a) Showing separation requirements expressed in time.



(b) Translating spacing requirements to demand. The contribution of AC2 has been highlighted. The crest of the coloured area provides an indication of instantaneous demand.



(c) Showing demand and capacity. The excess in demand imposed by AC2 has been highlighted.

Figure 4.5: Basic time line extended to show demand and capacity.

This visualisation forms a presentation of the *scheduling* generalised function and the *safety* and *capacity* objectives, as shown in Figure 4.7. Effectively, peaks in demand show the equivalence of underspacing and a capacity shortage. Doing so makes the need for action based on the current schedule clear immediately.

4.5.3 Delay and efficiency

The excess in demand is addressed by adjusting either the capacity or the schedule. The amount of excess demand determines the delay of the last aircraft in a sequence. The maximum delay can be a driver in a sequence manager's decision to apply other strategies (See Chapter 2).

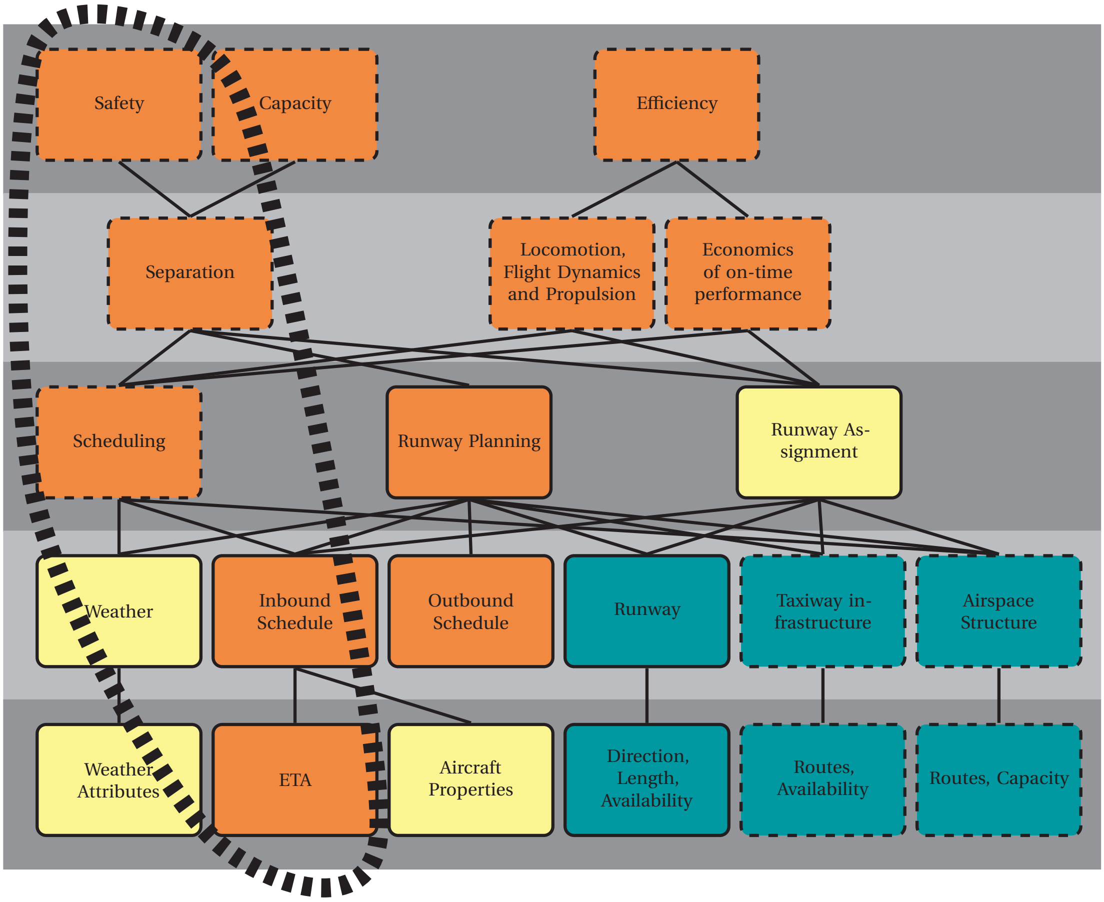


Figure 4.6: Visualisation of the spacing requirement on the abstraction hierarchy.

In current AMAN systems, this maximum delay is determined by planning all aircraft one after in the sequence that they are predicted to arrive. However, when arrival times are uncertain, the optimal sequence is not yet certain either. It is possible that the aircraft that is predicted to arrive as the last one in a sequence is not the last aircraft in the final optimal sequence. If that last aircraft is delayed to accommodate all predicted excess demand, the resulting solution will likely contain unnecessary deviation.

The new time line visualises demand (the blocks with a length equal to the spacing interval) and capacity (the area under the line) as an area. If a delay is applied as a solution strategy, a conflict resolves when the runway has provided available capacity equal to the excess demand. By visualising the demand as an area, it is possible to provide an a priori view of the maximum delay. This approach should give better support to the operator in selecting an appropriate strategy (see Figure 4.8).

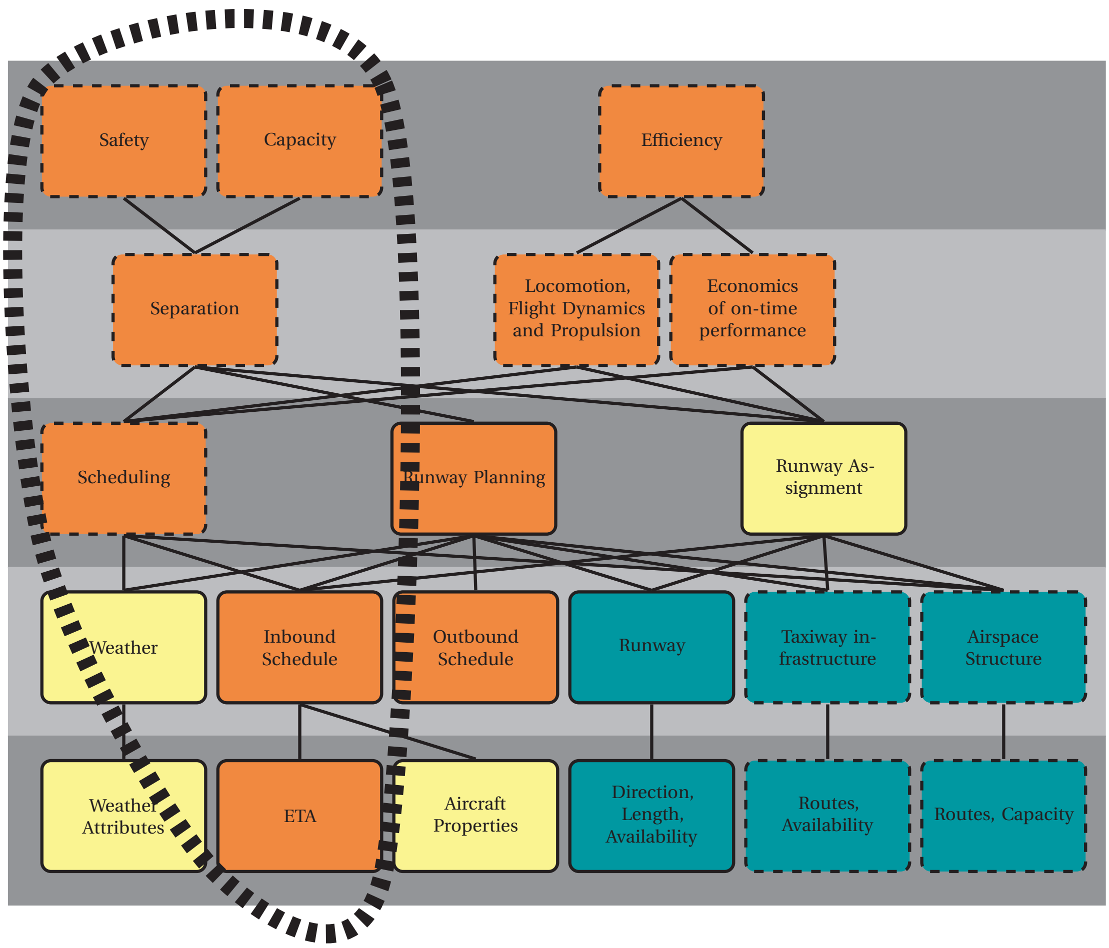


Figure 4.7: Visualisation of the demand/capacity balance on the abstraction hierarchy.

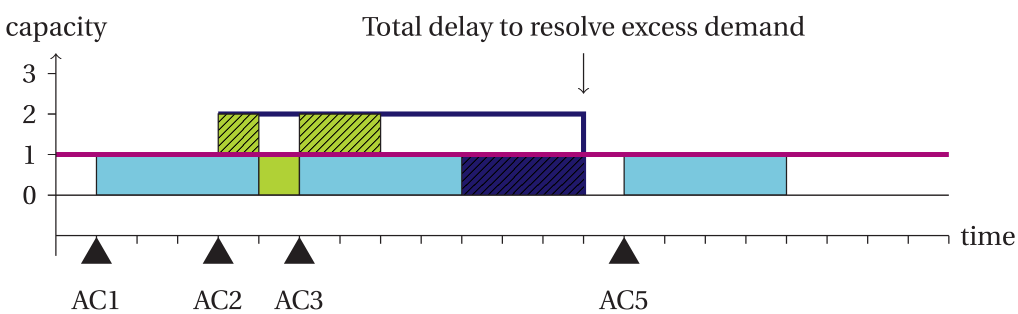


Figure 4.8: Visually relating excess demand to delay.

By using the a priori delay indicator, the new time line indicates the on-time performance of the predicted schedule. While it does not yet evaluate other strategies, such as having aircraft arrive earlier or opening an additional runway, it does give an initial indication of the efficiency of the current predicted solution (see Figure 4.9).

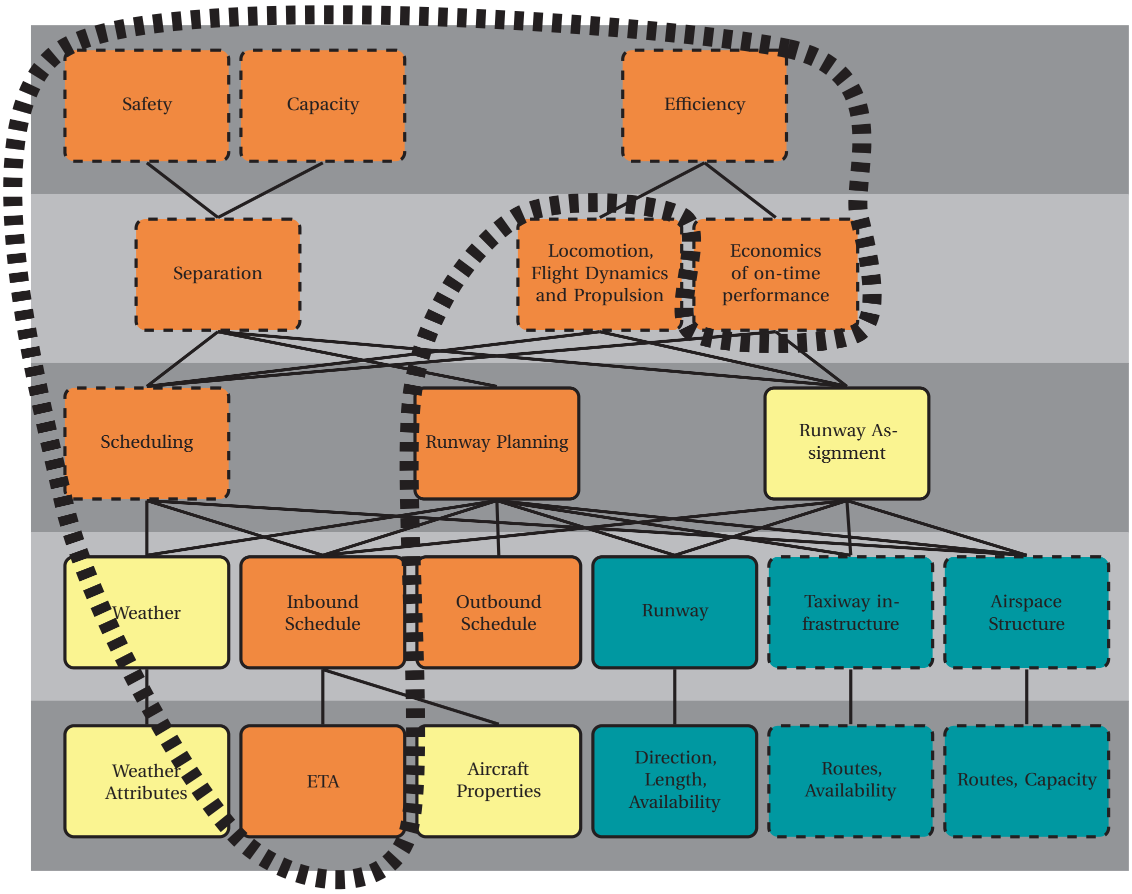


Figure 4.9: Visualisation of a priori delay on the abstraction hierarchy.

4.5.4 Integration: re-introducing uncertainty

Now that the time line has been extended to show the functional relations between the ETA and the purposes of the AMAN system, uncertainty can be reintroduced. To do so, the PDFs shown in Figure 4.3(b) are propagated to the occupancy expectation (i.e., the blocks representing demand).

Using a PDF for the ETA implies that the nature of the occupancy expectation blocks from Section 4.5.2 changes as well. When an aircraft is predicted to have a given probability of arriving at a given time, it has an equal probability of occupying the resource from that time for the duration of the separation interval.

The instantaneous expectation for runway occupancy due to the expected arrival at this time can be expressed as:

$$O_{P_i(t)}[t, t+s] = P_i(t) \tag{4.1}$$

In which $O_{P_i(t)}$ is the occupancy expectation for aircraft i at time t as a consequence of its probability of arrival $P_i(t)$. The s is the applicable spacing time interval.

The expectation for the demand at a given time becomes the integral of the arrival time probability for the spacing interval before it. This is demonstrated in Figure 4.10, and expressed mathematically as:

$$O_i(t) = \int_{u=t-s}^{t} P_i(u)\, du \tag{4.2}$$

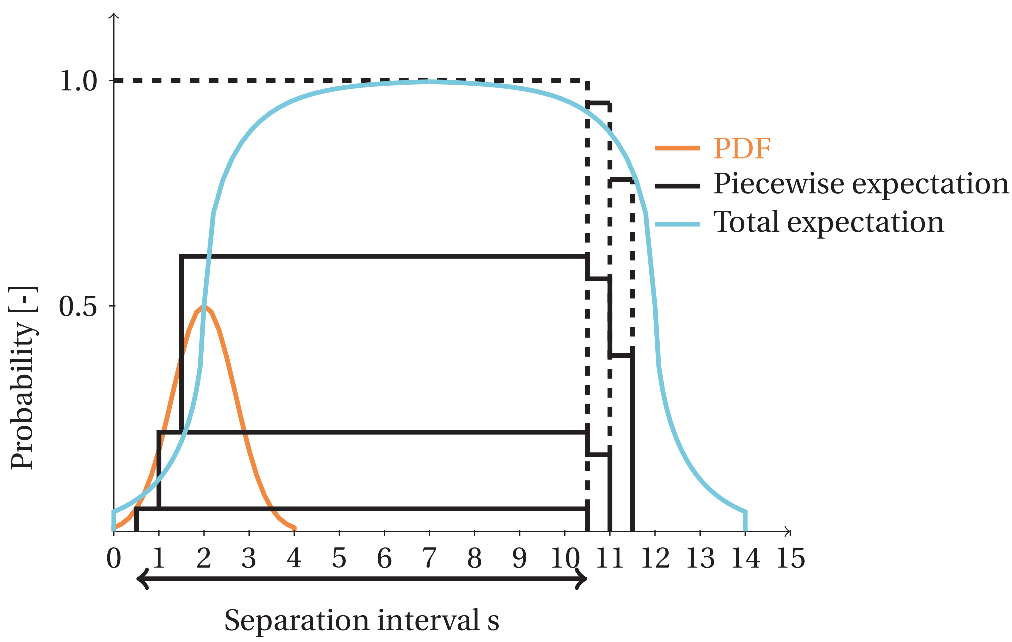


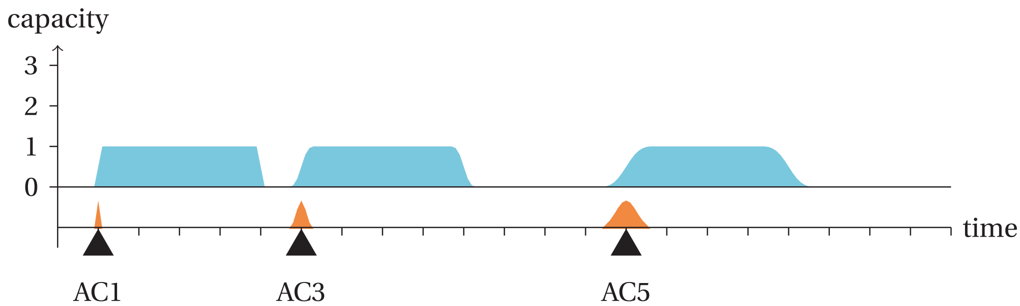
Figure 4.10: Calculation of the expectation value for occupancy. The PDF has been exaggerated for clarity.

Similar to the display without the uncertainty, the sum of the occupancy expectation for n aircraft generates the total expected occupancy:

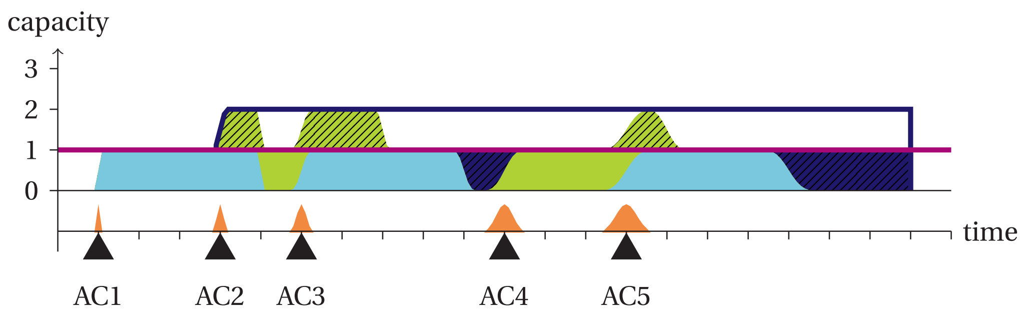
$$O(t) = \sum_{i=1}^{n} \int_{u=t-s}^{t} P_i(u)\, du \tag{4.3}$$

The resulting occupancy expectation can be presented on the time line (See Figure 4.11). The display still conveys the same principle: a higher occupancy expectation indicates a predicted excess. Note that an occupancy expectation above the capacity is not equivalent to a guaranteed excess as arrival times may still deviate, and the PDFs may move away from or toward each other.

Figure 4.11(b) shows that the a priori delay indicator still applies. This indicator is of particular value when demand is presented as an expectation value. The requested capacity remains equal regardless of how the schedule evolves due to prediction error. Therefore, the end of the delay indicator shows that a runway can provide the required capacity for the inbound aircraft.



(a) Showing the PDF of the arrival time and the resulting expectation value for resource occupancy.



(b) Presentation of expected excess demand, and the time required to resolve these problems by delaying aircraft. The contributions in demand due to aircraft AC2 and AC4 have been highlighted for clarity.

Figure 4.11: Integrated time line display.

By showing the a priori maximum delay, the operator can now see both the expected shortage of capacity and the expected consequences for the schedule without requiring planning first. The presentation provides the relations between the abstract functions *safety*, *capacity*, and *efficiency*, the generalized function *scheduling*, and the ETAs of different aircraft while allowing the ETA to be uncertain (see Figure 4.9).

4.6 Evaluating the new display: First experiment

Section 4.5 proposes a new version of the AMAN time line display. This display aims to support planning inbound aircraft to a single runway in the presence of uncertainty. An exploratory evaluation of the display was performed to evaluate the additions to the time line and to develop a suitable testing method [1].

This initial exploration found that the delay indicator reduced the number of control actions the experiment participants took. The visualisation of uncertainty did not noticeably change performance, however. Analysis of the subjective workload indicated a considerable learning effect.

This section will describe two follow-up human-in-the-loop experiments. The first experiment evaluated both the visualisation of uncertainty and the delay indicator. To further explore the findings of the first experiment, the second experiment focused solely on the visualisation of uncertainty.

The time line is suitable for a single-runway scenario. When multiple runways are available, the current display may indicate the total airport capacity but will not provide support in assigning aircraft to different runways or specific landing slots. Chapter 5 extends the display to a scenario with two runways and describes a similar experiment for that scenario.

4.6.1 Experiment method

Participants

While designed for Air Traffic Management (ATM) purposes, the concept display no longer provides information specific to aviation. Instead, it provides a display for any logistics planning process in which the time of arrival of an item or task is uncertain. Secondly, the display concept is designed for future scenarios in which arrival planning is to be performed up to two hours before the arrival. Currently, none such systems are operational and very few current ATCOs have experience with arrival planning at such long horizons. Finally, the number of operational sequence managers is highly limited in any case, making it difficult to find a sufficient number of qualified ATCOs for the task.

A first experiment evaluated the concept of visualisation in general. The lack of specifics for aviation allows participants with limited knowledge of the ATM process, and allowed to focus on the properties of the display itself.

A total of 16 participants took part in the first experiment. All were staff, students or former students from the Faculty of Aerospace Engineering at the Delft University of Technology. Five participants had prior knowledge of the spacing problem, and four had been participants in the earlier initial experiment on the same display.

Display

The experiment used a PC-based simulated AMAN. Participants were tasked with spacing aircraft to arrive at a single runway on a right-to-left moving time line (see Figure 4.12). Each aircraft was presented as a triangle on the time line at its predicted ETA and subject to prediction error. Participants could move aircraft along the time line by selecting the triangle using the mouse and dragging the aircraft to the preferred arrival time. This action would result in a hypothetical (or *probed*) solution.

Once aircraft were at the desired arrival times, the user could confirm these times by clicking a button on the screen, resulting in a desired arrival time[1] for the aircraft. If the probed solution was unsuitable, the user could cancel the probe using a different button. By cancelling a probe, all probed aircraft would return to the RTA that was assigned before that probe.

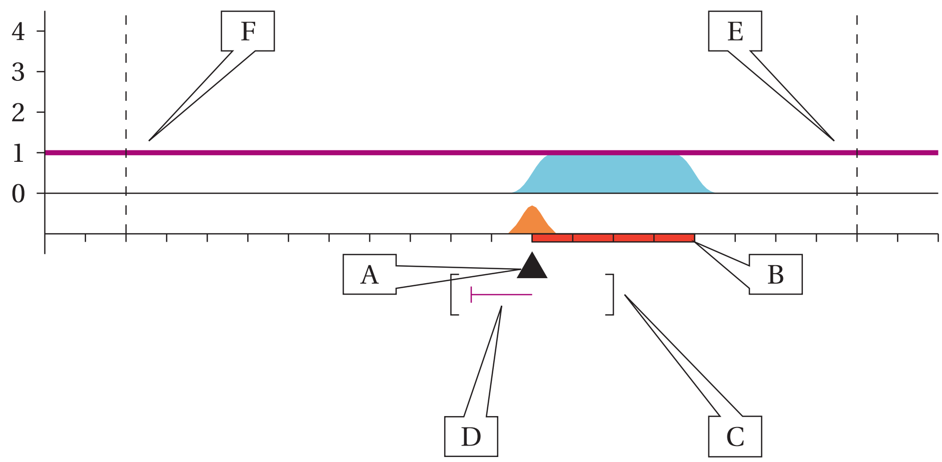


Figure 4.12: Symbols used in the experiment. A: The aircraft symbol at the predicted ETA (can be clicked and dragged using the mouse), B: The required spacing from the predicted ETA, C: The minimum and maximum RTA that can be requested, D: The current deviation from the optimal ETA, E: The earliest time at which aircraft RTAs can be requested, F: The latest time at which RTAs can be requested.

Besides the visualisation of uncertainty and the delay indicator, the time line display also showed several basic symbols relating to the characteristics of the selected aircraft. Figure 4.12 explains these different symbols.

To simulate the bounded control space, aircraft could only deviate a limited time from their initial ETA (C in Figure 4.12). The amount of possible deviation started at 20 minutes—earlier or later—and decreased linearly with remaining time to fly. This reducing control space introduced a part of the constraints defined

[1]From this point onward, the selected arrival times will be referred to as the Required Time of Arrival (RTA)

by *locomotion, flight dynamics and propulsion* in the abstraction hierarchy and is representative of the limitations in speed and endurance of aircraft.

Solving large conflicts with multiple aircraft by delaying all consecutive aircraft increases the delay required for downstream aircraft. The cumulative delay sometimes leads to too large required delays for individual aircraft. The delay indicator only provided the hypothetical solution using delay. The delay indicator turned red when delay would no longer be an option due to a lack of control space.

Since one of the objectives of AMAN is to minimise deviation from the optimal schedule, the display showed an indication of the deviation from the optimal ETA (D in Figure 4.12). The indicated optimal time was subject to the same prediction error as the ETA.

In the initial exploration, participants indicated that visualising the spacing requirement using occupancy expectation (see Section **??**) made it difficult to recognise the required spacing. As additional support, the minimum separation interval was shown as a red bar trailing each aircraft symbol (B in Figure 4.12).

The resulting experiment display is shown in Figure 4.13. The experiment presented the time line on a 24” LCD monitor.

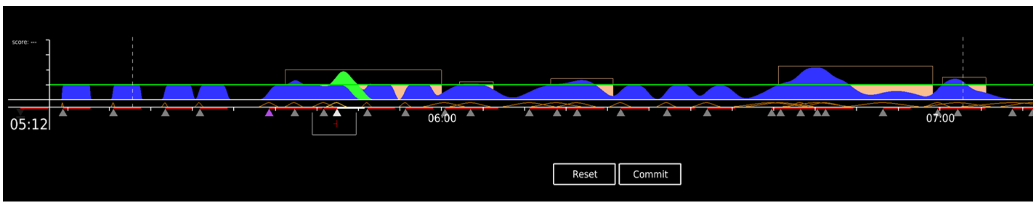


Figure 4.13: Experiment display as used during the experiment.

Simulation

The aircraft required a constant minimum spacing interval of 200 seconds. The time line had a horizon of two hours. Modifying the arrival time was only possible after an aircraft had appeared on the screen for ten simulated minutes to force participants to observe potential conflicts with trailing aircraft before taking action. Similarly, at ten minutes before arrival, modification of arrival time was disabled to simulate the impossibility of adjusting spacing during final approach. The display showed these limits as vertical lines. (see E and F in Figure 4.12).

In the current operation, the task of the sequence manager is often part of a broader role. For example, the approach supervisor operates the AMAN system at Air Traffic Control the Netherlands (LVNL). In a single-runway environment, the task demand for the operator, whose only task is to plan arrival times, would be too low to require fast decision-making and instead allow an extensive mental evaluation of options. The display aims to support decision-making. Most differences in results are expected when insufficient time is available for such an elaborate evaluation. Furthermore, such an experiment would require a very long time per participant to collect sufficient data points. To increase the task load and maximise the number of data points, the scenarios were played at 30 times real-time. Individual aircraft would thus be on display for four minutes actual time.

Traffic

The traffic schedule was defined beforehand to achieve the following properties:

- A comparable sample: Each participant would work on the same schedule;
- A feasible arrival schedule: Every spacing conflict was solvable if actions were taken in advance;
- Comparable uncertainty: The prediction error for a specific flight in every run was identical.

By building a scenario from a landing schedule and then superimposing a prediction error, traffic behaviour has been predefined. The following paragraphs describe the schedule-generating process.

Landing schedules were randomly generated to have a spacing according to a—forced—normal distribution with a running mean of 280 seconds over ten aircraft, and a standard deviation of 150 seconds. This distribution provided a mean buffer of 80 seconds between each landing interval and ensured that sufficient spacing would be available between local bunches to resolve these bunches.

Uncertainty consisted of a normal distribution with a linearly decreasing standard deviation as aircraft approached the runway. The initial standard deviation was randomly assigned between 50 and 200 seconds. The standard deviation at landing was set to 0, so uncertainty would decrease to none at the end of the time line.

The prediction error was defined as a cumulative probability within the aircraft's uncertainty distribution. The cumulative probability was randomly assigned between 0.1 and 0.9. The combination of uncertainty and prediction error resulted in a varying error per aircraft that would decrease over time. At the same time, the prediction error would always be consistent with the indicated uncertainty. The distribution of uncertainty and prediction error were distributed evenly over all flights.

To achieve a comparable level of difficulty between scenarios, samples were set up to require a comparable amount of effort to resolve. A set of samples was randomly generated according to the described process. Subsequently, the optimal solution (i.e., minimal total deviation) was approximated by spreading out—both through advancing and delaying—local bunches until all aircraft were spaced with a buffer of at least 20 seconds. The required actions were then verified to be achievable within the maximum control space of each aircraft. Unachievable scenarios were discarded, and the required number of actions and their magnitudes were compared. Finally, of the comparable samples, a selection was made based on inspection of the amount and size of conflicts at the two-hour time horizon which included the maximum prediction error.

The four selected scenarios consist of two pairs of comparably complex scenarios. Scenarios 1 and 4 require a relatively high amount of change for the optimal solution (44 and 48 minutes total change, respectively). Scenarios 2 and 3 have a somewhat lower amount of required change, which suggests a less complex spacing task (22 and 18 minutes, respectively).

Procedure

As mentioned before, an initial experiment using the display demonstrated that considerable training was necessary before participants were able to use the information provided in the display properly [1]. To reduce training effects in this experiment, participants were trained in two phases: First, a step-by-step walk-through in which all display concepts (shown in Figure 4.12) were introduced one at a time. Secondly, participants would train the use of a specific display in four training runs before performing the experiment runs for that display type.

The preliminary experiment also showed effects due to an apparent lack of interest in the final result of the planning, leading to a rather conservative application of buffers. To motivate participants in providing sufficient spacing, while at the same time also minimising deviation and actions, a running score was provided simultaneously with a fictitious but realistic high score. The score was based on the combination of spacing error, the amount of deviation, and the number and timing of deviation actions. The score was weighted to stimulate participants to avoid underspacing (i.e., conflicts between aircraft at landing).

4.6.2 Independent variables

The experiment tests two elements of the display: the presentation of uncertainty and the availability of the delay indicator. Therefore, the two independent variables are the presentation of uncertainty (U/N) and the presentation of the total required delay (D/N). The first variable was tested within-subject. However, the second variable was tested between groups due to time constraints. Participants would only see a particular scenario either with or without the delay indicator. Assignment of displays and scenarios was based on a latin square to counterbalance training effects (see Table 4.1).

Table 4.1: Latin square assignment of scenarios and displays to participants.

Part.	Experiment run															
	1		2		3		4		5		6		7		8	
	S	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D
1	1	2	2	1	4	2	3	1	1	3	2	4	4	3	3	4
2	2	4	3	3	1	4	4	3	2	1	3	2	1	1	4	2
3	3	1	4	2	2	1	1	2	3	4	4	3	2	4	1	3
4	4	3	1	4	3	3	2	4	4	2	1	1	3	2	2	1
5	1	2	2	1	4	2	3	1	1	3	2	4	4	3	3	4
6	2	4	3	3	1	4	4	3	2	1	3	2	1	1	4	2
7	3	1	4	2	2	1	1	2	3	4	4	3	2	4	1	3
8	4	3	1	4	3	3	2	4	4	2	1	1	3	2	2	1
9	1	2	2	1	4	2	3	1	1	3	2	4	4	3	3	4
10	2	4	3	3	1	4	4	3	2	1	3	2	1	1	4	2
11	3	1	4	2	2	1	1	2	3	4	4	3	2	4	1	3
12	4	3	1	4	3	3	2	4	4	2	1	1	3	2	2	1
13	1	2	2	1	4	2	3	1	1	3	2	4	4	3	3	4
14	2	4	3	3	1	4	4	3	2	1	3	2	1	1	4	2
15	3	1	4	2	2	1	1	2	3	4	4	3	2	4	1	3
16	4	3	1	4	3	3	2	4	4	2	1	1	3	2	2	1

S = Scenario, D = Display type
Display types: 1 = no uncertainty, no delay indicator (N/N), 2 = no uncertainty, delay indicator (N/D), 3 = uncertainty, no delay indicator (U/N), 4 = uncertainty, delay indicator (U/D).

As explained in Section 4.6.1, some effort was spent to ensure that the scenarios were of comparable complexity. As these scenarios ended up being still somewhat different, they potentially formed a third independent within-participants variable.

The final independent variable in the experiment is the participants. The analysis will be performed within-subject as much as possible. However, some between-participants analysis may be worthwhile in explaining the results.

4.6.3 Dependent measures

To determine the operator's decision-making performance, the quality of the schedule and the way of achieving that schedule is determined. These are quantified by two parameters, measured at the point that the aircraft leave the time line:

- The amount of underspacing between landing aircraft, and
- The total absolute amount of deviation from the initial schedule larger than the total absolute amount required for an optimal schedule. By selecting total absolute deviation, the metric is not sensitive to the selected solution nor the number of changes in RTAs, as long as the total deviation is minimised.

The main benefit of a long time horizon in arrival management lies in the ability to act early, resulting in small changes in speed. To determine the success in achieving early actions, the Estimated Time Enroute (ETE)[2] of an aircraft at each confirmed RTA change was recorded.

Furthermore, the subjective workload was measured using Instantaneous Self Assessment (ISA) [11]. Participants were prompted using an on-screen continuous analogue scale every 30 seconds. They were alerted to respond using an audible signal over headphones. Before the experiment, participants were instructed to respond to each ISA probe. However, they were always able, and allowed, to continue using the AMAN display during an ISA probe.

Finally, participants were interviewed after the complete experiment to elicit both their understanding and use of the display and their opinion on the added items. This interview was structured to discuss all display attributes twice. In the first pass, participants were only asked to explain what the display attributes indicated and how they worked. During the second pass, participants were asked to explain whether and, if so, how they would use the information.

4.6.4 Hypothesis

The purpose of the display is to improve the planning process. An optimal planning can be defined as one that results in as little change to the individual flights' schedules while maintaining an equal or better spacing. Furthermore, any changes should be requested as early as possible to minimise the aircraft's fuel consumption. The cost of planning can then be defined by the total amount of change requested during the planning process (including corrections due to uncertainty) and the remaining flight time in which to achieve that change.

Based on this definition, three hypotheses are tested:

> $H1$: When uncertainty information is provided, the total amount of excess deviation is decreased, and the changes in RTA are given earlier in the flights, both while not compromising on minimum spacing.

[2] The ETE is the predicted remaining time to fly (i.e., the difference between the current time and the predicted ETA. On the time line, the ETE is the distance between the aircraft symbol and the end of the time line.

$H2$: When the delay indicator is provided, the total amount of excess deviation is decreased, and RTAs are given earlier in the flights, both while not compromising on minimum spacing.

$H3$: When the delay indicator is provided, the amount of underspacing decreases.

4.7 First experiment results

Of the 128 measurement runs, two runs, with two different participants, had to be discarded due to problems with the recording software. One other experiment run was stopped and restarted shortly after the beginning. Otherwise, no other issues occurred that may have disturbed the measurements. The following subsections will describe the objective results, the results of the interviews, and the conclusions drawn from the first experiment.

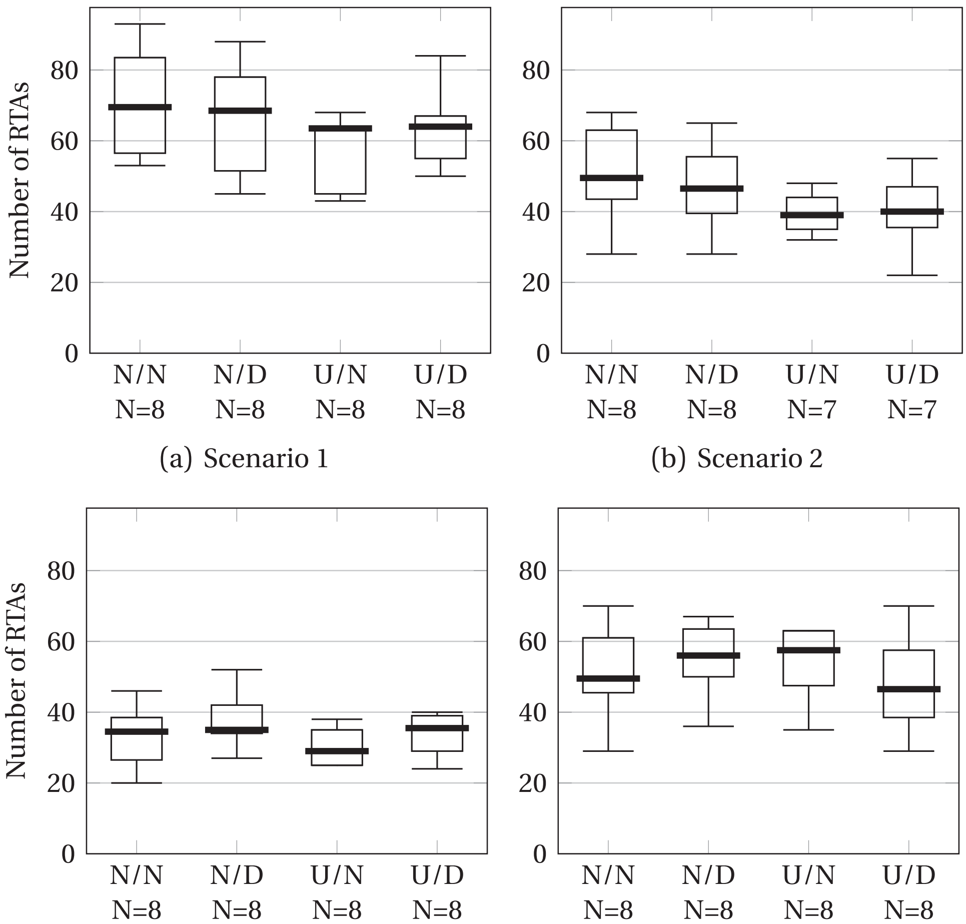
4.7.1 Objective metrics

Section 4.6.1 explains that considerable differences exist between the four scenarios in the amount of correction required for an optimal solution. Figure 4.14 reflects this in the total number of actions that were performed by all participants; As expected, less RTA changes were applied in Scenarios 2 and 3 as compared to Scenarios 1 and 4.

Since Scenarios 2 and 3 require far less deviation than Scenarios 1 and 4, results from these scenarios cannot be compared directly. To enable comparison of results over the different scenarios, spacing errors and corrective actions have been normalised by dividing them by the total required change for the optimal solution. However, the results will continue to be analysed per scenario to account for the large differences between scenarios.

Each participant has only evaluated each scenario either with or without the use of the delay indicator. The use of the delay indicator is thus coupled to each scenario. Therefore, subject, scenario and delay indicator are not independent when evaluated together. By assessing the results per scenario, the experiment no longer supports a statistically valid within-subjects analysis of the effect of the delay indicator. The latter will, therefore, only be discussed in general terms noting that a specific study is required to reach proper conclusions.

Differences were tested for significance using the Wilcoxon Signed Rank test for paired measurements and the Wilcoxon Ranked Sum test for independent samples. Since the cases with and without the delay indicator are evaluated separately, all tests will apply a Bonferroni correction of 2 to a criterion for statistical significance (p) of 0.05.



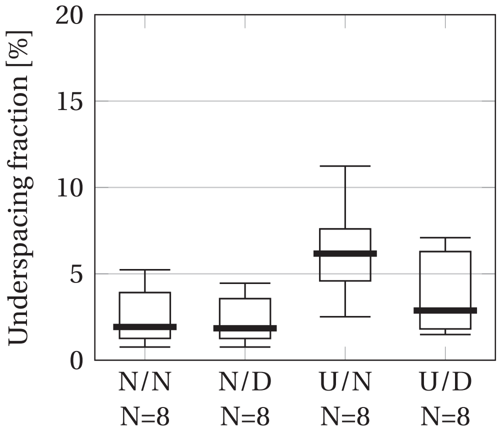
(a) Scenario 1 (b) Scenario 2

(c) Scenario 3 (d) Scenario 4

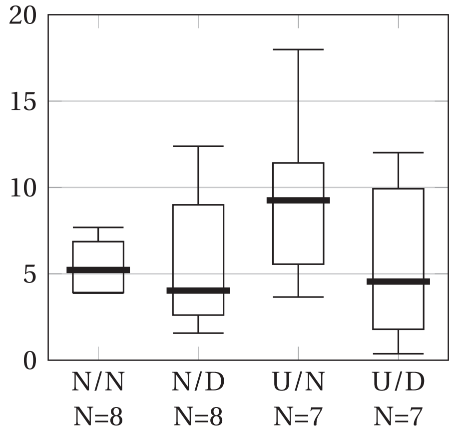
Figure 4.14: Number of RTAs requested.

When uncertainty indicators are available, the participants seem to apply a smaller number of corrective actions in Scenarios 1 and 2 (see Figure 4.14). The differences in the number of corrective actions in Scenarios 3 and 4 are less pronounced.

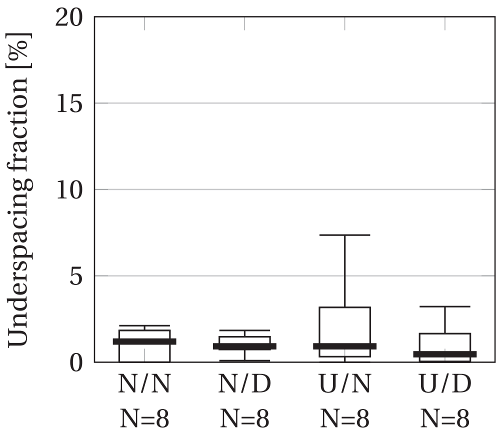
The primary objective in using the time line display is to ensure no less than minimum spacing at landing. Figure 4.15 shows the total sum of underspacing between consecutive aircraft. In the display, this underspacing is represented as the amount overlap of the occupancy indicators at landing. The improved performance that was hypothesised for this experiment was under the condition that the minimum spacing is not compromised. When the delay indicator is not available, the addition of the uncertainty indicator tends to increase the underspacing when compared to the situation without the uncertainty indicator ($z = 2.59, p < 0.01, r = 0.33$). When the delay indicator is available, the addition of the uncertainty has no effect ($z = 1.67, p = 0.05, r = 0.21$).



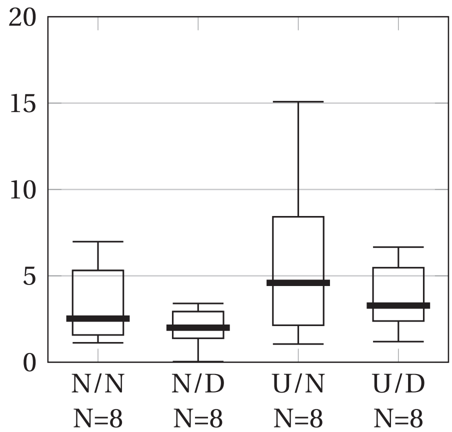
(a) Scenario 1



(b) Scenario 2



(c) Scenario 3

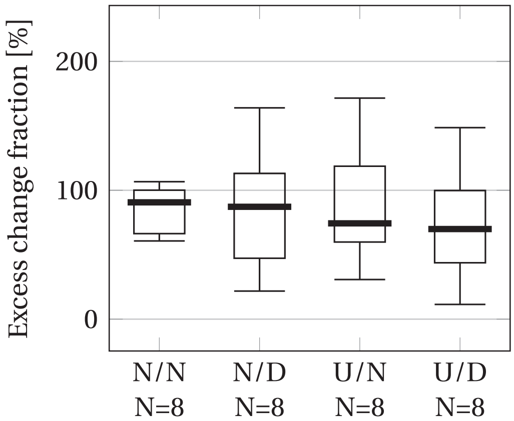


(d) Scenario 4

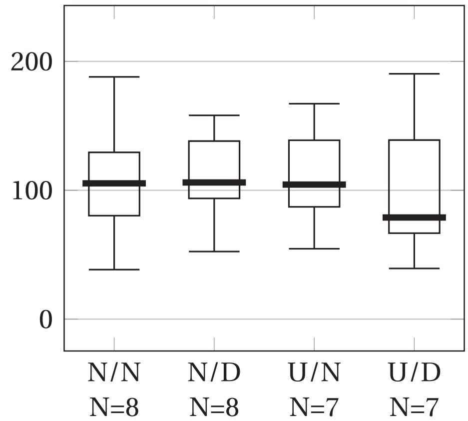
Figure 4.15: Total insufficient spacing at landing as a fraction of the total expected spacing for an optimal solution.

When spacing is sufficient, the optimal solution is that for which minimal deviation has been applied (see Figure 4.16). From the graphs, a first observation is that the total amount of change is often more than twice the required amount (100% excess). The addition of uncertainty to the display without the delay indicator tends to decrease the excess spacing ($z = 2.58, p < 0.01, r = 0.33$). The result is not significant when the delay indicator is provided ($z = 1.99, p = 0.05, r = -0.35$). Displaying the total delay appears to have no effect.

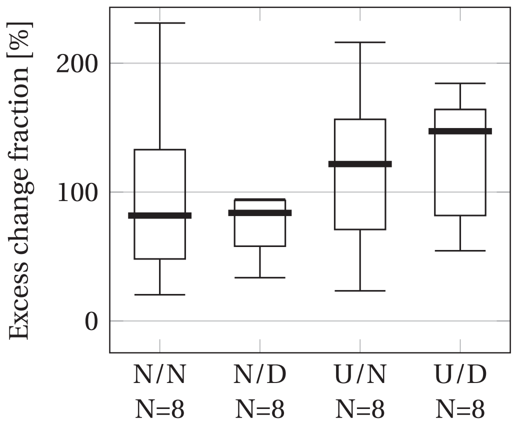
After the experiment, several participants pointed out that the visualisation with uncertainty made it more difficult to recognise overlap immediately after aircraft appeared on the right end of the time line. The occupancy expectation for a given aircraft appeared ten—simulated—minutes after the aircraft symbol appeared to force evaluation of potential trailing aircraft (see Section 4.6.1).



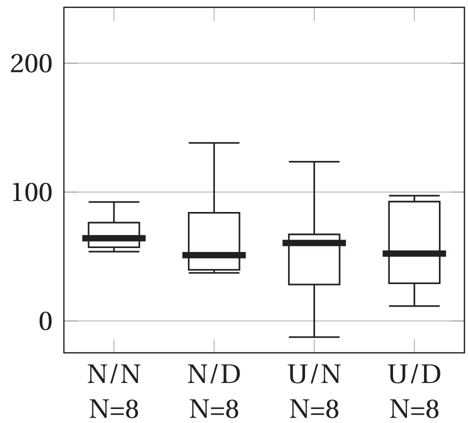
(a) Scenario 1



(b) Scenario 2



(c) Scenario 3



(d) Scenario 4

Figure 4.16: Amount of requested deviation more than required for optimal solution. Note that the indicated amount is the excess amount of spacing above the required amount for the optimal solution (i.e. 100% is twice the amount of deviation than required for the optimal solution).

The display without uncertainty indication gives a powerful indication of conflicting aircraft since all occupation expectations stack vertically. These form quickly rising towers when a spacing conflict is predicted. In contrast, the uncertainty spreads out the occupation expectation, leading to a lower total occupancy expectation and, therefore, less noticeable potential underspacing.

To further analyse the relationship between remaining time to fly (ETE) and selected strategy, heat maps in Figure 4.17 show which combination of ETE and deviation had a preference. A red colour indicates that many flights—over all runs for that scenario/display combination for all participants combined—were issued a similar change in ETA at a similar ETE. The white line provides a histogram of the number of RTAs changes at different ETEs.

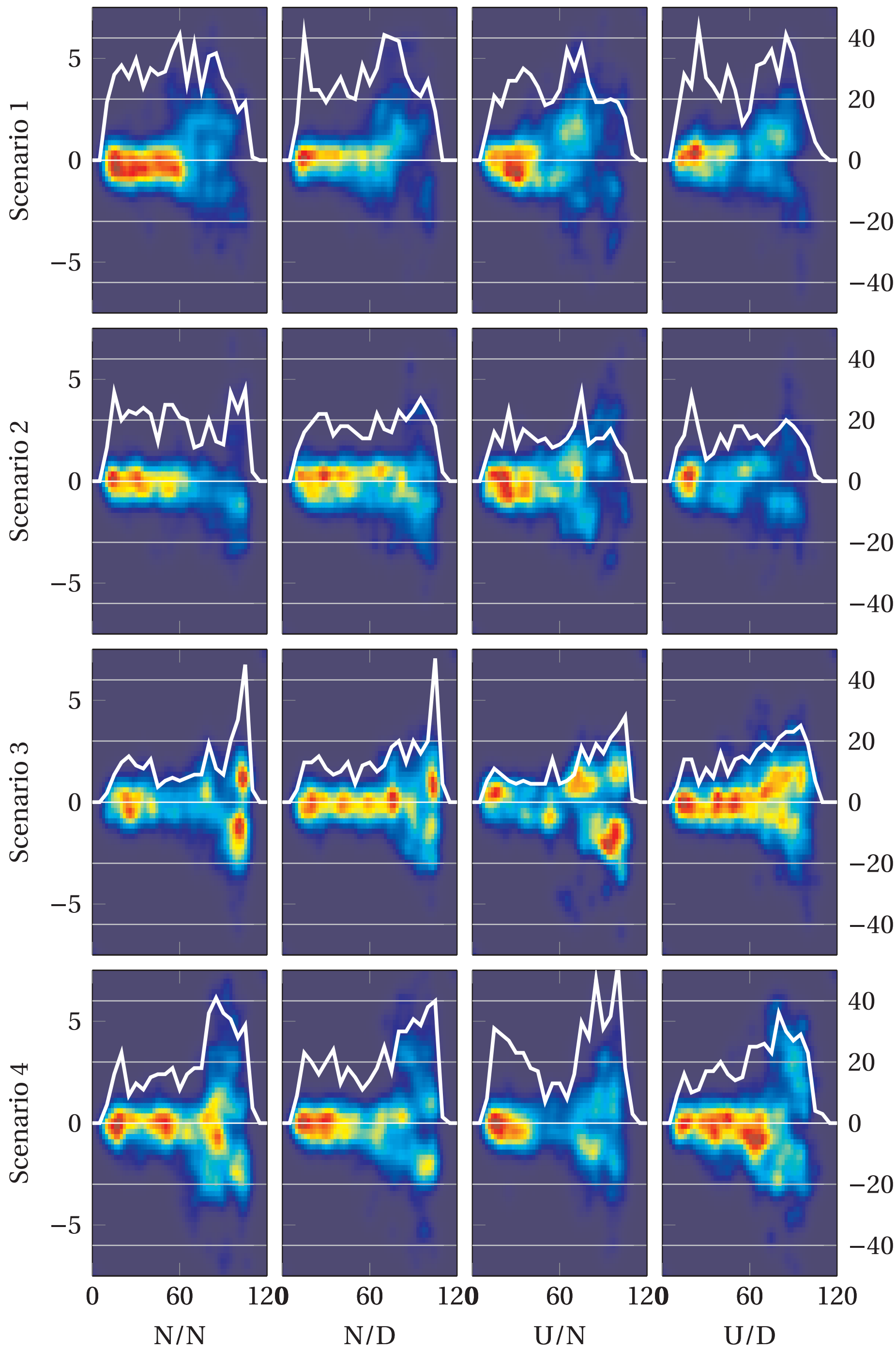
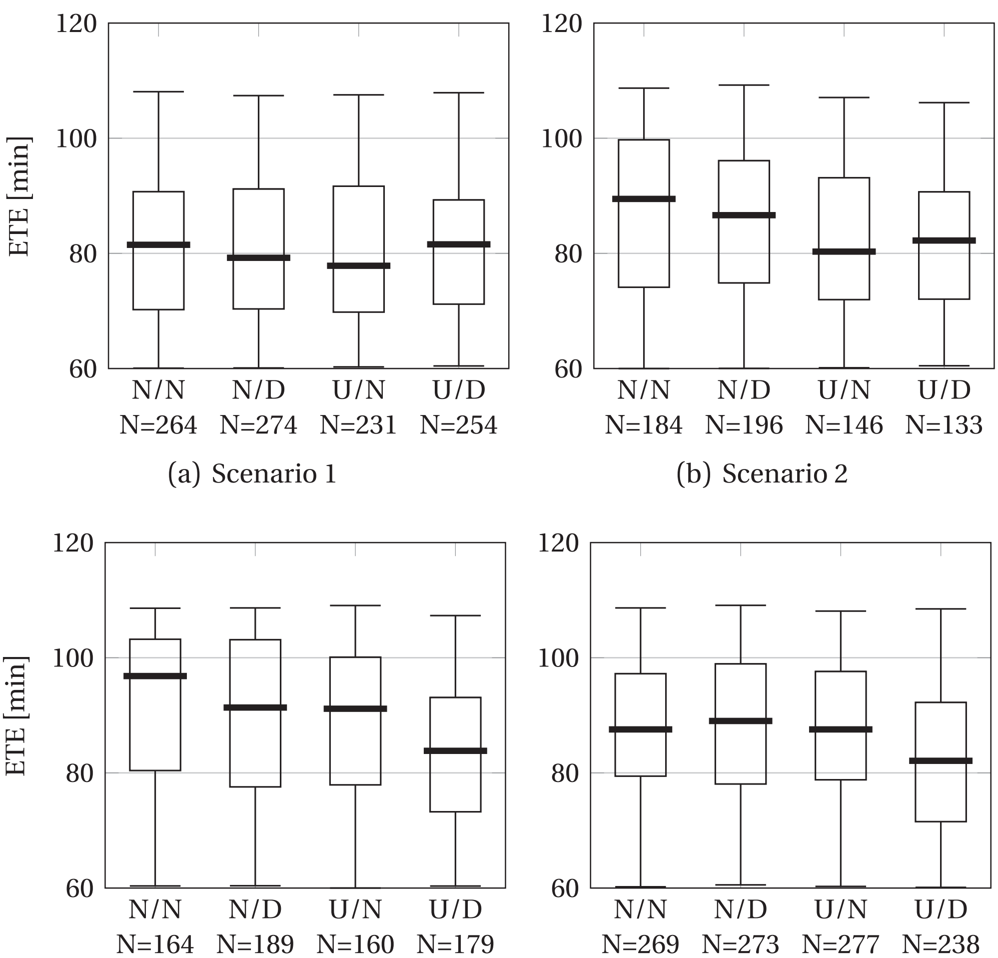


Figure 4.17: ETE and requested change of RTAs. The heatmap indicates the ETE (x-axis, [min]) and amount of requested change (left y-axis [min]). The white plot indicates the number of RTAs per 5 minutes (right y-axis).

Figure 4.17 shows that participants tended to respond by immediately unstacking these aircraft as soon as they appeared on the display. This is especially noticeable for Scenario 3. Such immediate action suggests that the indication of predicted conflicts was deemed more important than the uncertainty in the ETA of the involved aircraft.

The scenarios with high traffic density (1 and 4) often required many adjustments of the RTA short before arrival. In these cases, spacing errors reappeared, and any possible spacing buffers had been used up. Figure 4.17 shows that the participants applied a similar strategy in Scenario 1 for the display without delay indicator and without uncertainty indication. In this case, most conflicts appear to be resolved by small delays over the full left half of the time line. The heat maps show that this strategy is not applied when the delay indicator is provided on the display.



(a) Scenario 1 (b) Scenario 2

(c) Scenario 3 (d) Scenario 4

Figure 4.18: ETE at which RTA was requested.

Figure 4.18 aims to quantify the change in timing seen in Figure 4.17. This figure shows the distribution of the times at which an RTA change was confirmed during the first (right) 60 minutes of the time line. For the display without delay indicator, the presence of uncertainty information does not affect the preferred lookahead for corrections ($U = 1209992, p = 0.59, r = 0.01$). When the delay indicator is provided, the ETE at which RTAs are requested decreases slightly. While statistically significant, the effect is negligible ($U = 1145853, p < 0.01, r = 0.07$). Note that these tests are performed assuming that scenario and subject have no effect on these measured parameters. The box plots for Scenarios 2 and 3 suggest a trend for the time of the decision to be delayed more as more visual aids are presented. This suggests that the additions will only influence decisions when participants have more options in assigning RTAs.

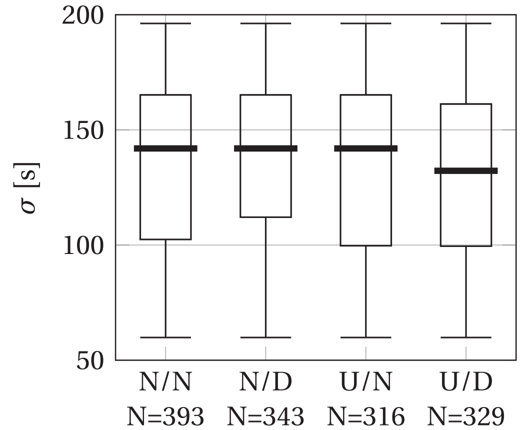
Figures 4.15 and 4.16 did not show significant deviations of the planning results. To see whether the decisions were influenced by the uncertainty of the aircraft, Figure 4.19 plots the distribution of the standard deviations of the initial uncertainty of the aircraft for which the RTA was changed. Since uncertainty decreases with the remaining time to fly, the initial uncertainty is a good indicator of a preference for higher or lower uncertainty at a given ETE.

As described earlier, a first difference between the display with and the display without uncertainty is that participants often spaced aircraft as soon as they appeared on screen in the display without uncertainty. In contrast, participants would wait longer in the display with uncertainty. The first 30 simulated minutes on the display are, therefore, excluded to avoid this effect of unstacking as soon as the aircraft appeared on the display.

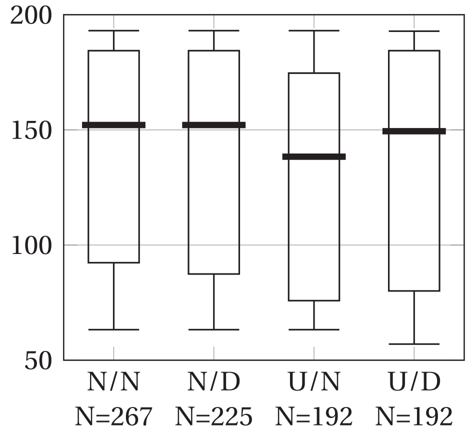
Figure 4.19 shows the distribution of the initial uncertainty of the flights that received RTA changes. Since the uncertainty decreases linearly with the remaining time to fly, the initial uncertainty is an indication of the relative uncertainty of a flight. If participants had structurally assigned more RTA changes to higher or lower uncertainty flights, the distribution should shift accordingly.

Figure 4.19 shows that participants had no clear preference for certain or uncertain aircraft in deciding which aircraft to adjust. There is no significant effect of the visualisation of uncertainty between the choice of adjusted aircraft when the delay indicator is not provided ($U = 1189764, p = 0.09, r = 0.03$). When the delay indicator is provided, a significant but irrelevant difference exists ($U = 1165463, p = 0.01, r = 0.05$).

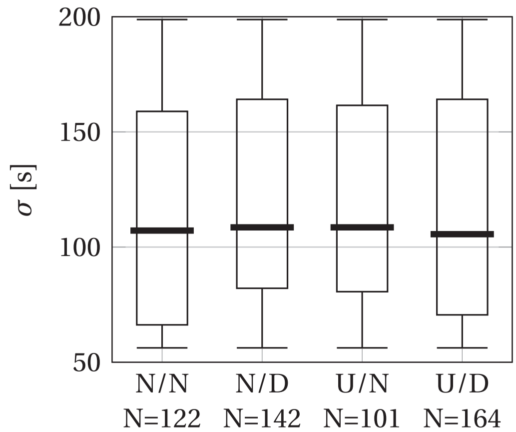
This could imply that the visualisation of uncertainty was not required, as participants did not include the amount of uncertainty of the individual flights in the decision-making processes. A confounding factor in this measurement is that the standard deviation of uncertainty of the flights is distributed continuously between 50 and 200 seconds (see Section 4.6.1), which reduces the distinction between the individual flights.



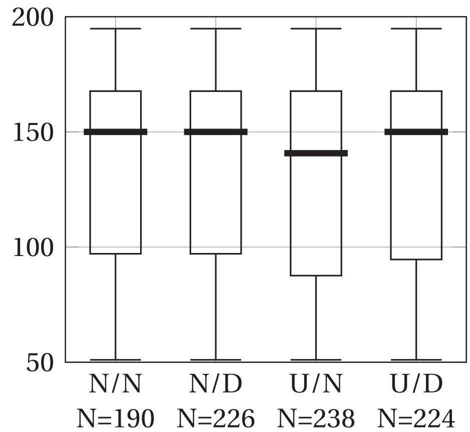
(a) Scenario 1



(b) Scenario 2



(c) Scenario 3



(d) Scenario 4

Figure 4.19: Standard deviation of initial uncertainty for a flight for which an RTA change was requested. Note: only aircraft at less than 90 minutes before landing were included.

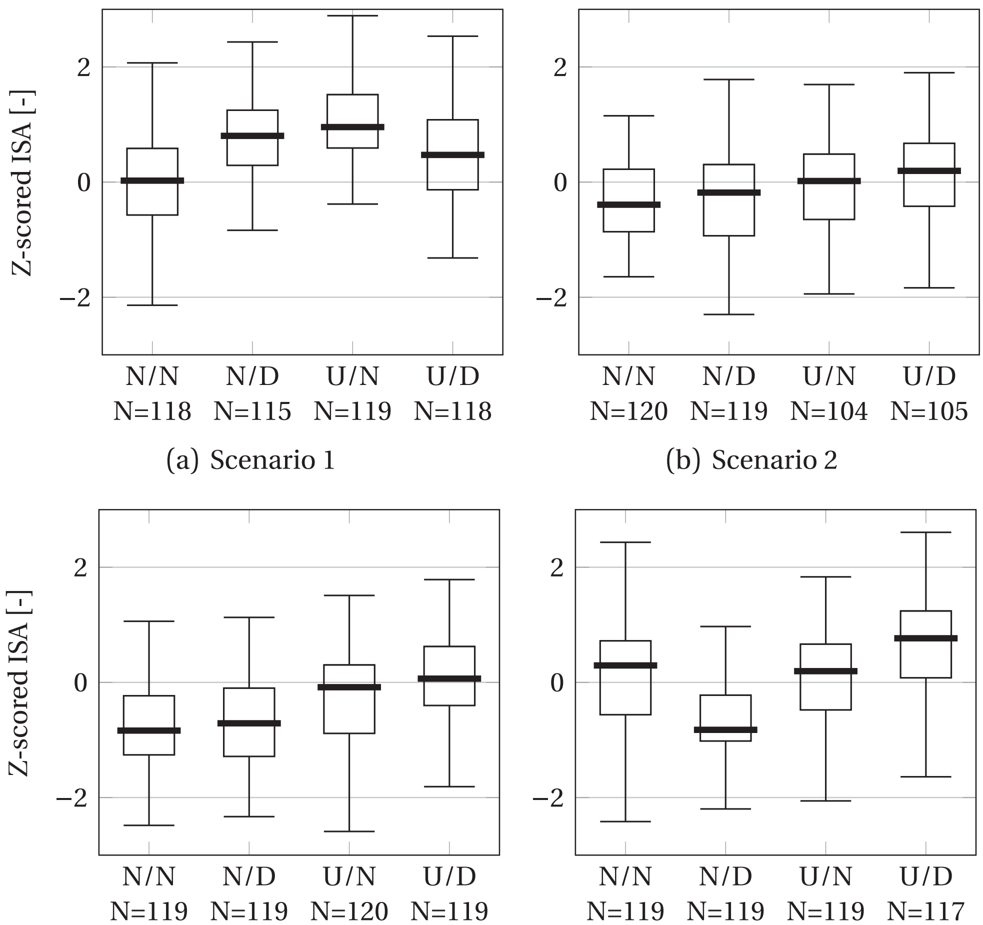
4.7.2 Subjective input

The objective results suggest that the addition of both new elements did not improve operator performance. This section first evaluates the perceived workload, then the participants understanding and use of the aircraft-specific symbols, followed by the uncertainty indication, and finally, the delay indicator.

Workload

Only a few ISA probes did not receive a response within the 30 seconds to the next probe (in total, 120 responses are expected per display, a lower number indicates missed experiment runs or missed probes). In general, the data suggest that the addition of new display components increased workload (see Figure 4.20). The visualisation of uncertainty lead to an increase in workload when the delay indicator is shown ($z = -7.91, p < 0.01, r = -0.27$) and also when the delay indicator is not shown ($z = -7.02, p < 0.01, r = -0.23$).

This finding corresponds with the verbal responses of the participants, most of whom indicated that understanding the extended, more complex display with uncertainty required more attention. Two exceptions are the addition of a delay indicator in Scenario 1 with uncertainty and Scenario 4 without uncertainty. In general, the addition of the uncertainty indicator also tends to increase the subjective workload.



(a) Scenario 1

(b) Scenario 2

(c) Scenario 3

(d) Scenario 4

Figure 4.20: ISA workload reported in experiment.

Aircraft symbols

- When aircraft symbols (the black triangle in Figure 4.12) below the time line would overlap (i.e., multiple aircraft are predicted to arrive at almost the same time), the display would vertically separate them to show the individual aircraft. Nine participants indicated that this vertical stacking was their main cue for initial spacing as soon as new aircraft appeared.

- At high traffic volumes, four participants indicated that they would only space the aircraft symbols evenly, regardless of the spacing indicator or the occupancy expectation. This behaviour was recognised by other participants as well and suggests that participants tended to fall back on the basic information at high task demand.

- All participants understood the meaning of the control space indicator and the fact that the control space would shrink as the aircraft approached the end of the time line. Six participants actively aimed to save some remaining control space for later corrections. However, these participants did not use the uncertainty information to decide on the amount of control space buffer.

- Since participants were scored both on the amount of underspacing and the amount of deviation from the predicted ETA, six participants actively tried to minimise the deviation when, due to prediction error, excess space became available. Two participants typically selected the aircraft with the least deviation to resolve a spacing conflict between two aircraft.

- Without uncertainty, the occupancy blocks were used to determine the required spacing. These functioned in the same manner as the spacing indicator but were a stronger visual attractor.

Uncertainty indicator

When provided with the uncertainty display, participants had widely varying strategies.

- Some participants focused on the PDF and separated the flights by ensuring that the tail of the PDF did not overlap the separation indicator. While this ensured that the uncertainty of the aircraft under control would never lead to a spacing error, the separation indicator was bound to the predicted arrival time of the leading aircraft, therefore not including that prediction's uncertainty. The other aircraft could thus still get closer to the aircraft under control and cause a conflict.

- Two participants used the width of the PDF to decide on setting an RTA. One participant sequenced aircraft with smaller uncertainties together to take benefit from the higher degree of certainty of the spacing between those aircraft. The other participant moved the most certain aircraft first as those could be moved to a definitive arrival time.

- Participants used the visualisation of the occupancy expectation in two distinct ways: One group would continuously aim to get all the excess below the maximum available capacity. This technique provides sufficient separation but leaves large amounts of excess spacing between aircraft and is often not feasible due to the high number of aircraft. The second group would try to spread out the excess initially and then further detail the spacing to achieve sufficient separation. While not leading to excess spacing, this technique takes more time and often leads to multiple consecutive actions on the same aircraft.

Delay indicator

In principle, the technique of roughly spreading out excess demand at the right side of the time line would benefit from the delay indicator. This indicator would immediately show whether adjustment of an aircraft would be sufficient to allow spacing of all aircraft involved in a conflict. However, only three participants appeared to have correctly understood the relation between the delay indicator and the excess demand.

- One participant did actively aim to create gaps between the end of the delay indicator and the occupancy expectation of the next bunch of aircraft. This strategy is equivalent to pre-planning buffers between groups of conflicts.

- Four participants recognised that a set of aircraft spanned by a single delay indicator formed a group of aircraft involved in a conflict. Resolution of such a conflict is possible without affecting aircraft in other groups.

- Since the delay indicator only showed how far back aircraft would have to move, two participants indicated that this visualisation generated a bias toward delaying aircraft instead of planning them to arrive earlier.

4.7.3 Conclusions on the first experiment

Despite the considerable training time, the responses indicate that only a few participants fully understood the objects on the diagram and their relation to solving the planning problem. The visualisation of uncertainty and a priori delay had a small effect on how participants scheduled the simulated aircraft. In most cases, visualisation of uncertainty led to delayed decisions without significant improvements to the quality of the resulting schedule. This effect is contrary to the objective of the display.

Only three participants recognised the equivalence between the area indicating excess runway occupancy and the delay indicator. Furthermore, only a few participants identified that creating gaps in the delay area would split conflicts into more manageable problems. Adding the delay indicator provided little help in solving the planning problem.

Some aspects of the experiment may have undermined the ability of participants to change their strategies based on the available information:

- In general, participants found the traffic volume to be high. With high traffic volume, limited options exist for solving the planning. The participants often employed a reactive strategy consisting of spacing all aircraft at regular intervals and continuously correcting errors as aircraft approach the end of the time line.

- The high task demand especially showed when participants would have probed a change but failed to commit that change since that would require clicking a button outside the time line. Several participants suggested using the keyboard for commands such as committing or cancelling a probed change.

- Participants were introduced to the display through a series of training scenarios. These scenarios introduced all aspects of the display. However, the lack of understanding by some participants suggests that training was insufficient. During discussions afterwards, some participants suggested that providing basic strategies during training would have helped them better apply the information available through the new display items.

- Figure 4.17 shows a high amount of small corrections in the last phase of the flight. These corrections would never change the sequence but only correct small overlaps. In those corrections, participants used little more information other than that the two aircraft were not sufficiently separated. The presence of uncertainty or delay visualisation did not affect the decision-making process during this phase of the flight.

4.8 Second experiment

As the first experiment failed to provide conclusive results, a smaller second experiment was performed to analyse several trends in the first experiment. The experiment design aimed to eliminate some of the confounding factors from the first experiment. This section will only discuss the differences from the previous experiment.

The most important confounding factor was the introduction of two display features under test: uncertainty and the delay indicator. Since this research focuses on uncertainty, the second experiment display always showed the delay indicator.

A second factor was the amount of traffic and the effort required to separate that traffic. The high task demand load prevented most participants from properly evaluating different solutions. Furthermore, participants spent much time correcting small spacing errors toward the end of the flight (the left side of the time line). Since these corrections only consisted of addressing underspacing as it appeared, they did not require participants to evaluate the available options other than moving aircraft apart until the conflict was resolved.

Based on the information provided by the first experiment, two changes were made to allow participants to consider different options: The traffic density was reduced considerably, and an automated algorithm would try to correct for small underspacing errors from 60 simulated minutes before landing.

The next subsections will first discuss the simulation and traffic samples to describe the new scenario under test. The following subsections will then discuss the hypotheses, the participants, and the experiment properties.

4.8.1 Simulation

The automated spacing algorithm applied a spacing technique based on first-arrive-first-served. This algorithm is also in operation with LVNL: Aircraft are initially planned on the time that they are expected to arrive. If the next aircraft is predicted to arrive too close to the previous one, it is delayed until sufficient spacing is achieved. The algorithm is neither allowed to make aircraft arrive earlier nor is it allowed to change the predicted sequence.

The automated algorithm in this experiment only operated on the last 60 simulated minutes—the left half of the time line. Furthermore, the algorithm is not allowed to move aircraft beyond their available control space. Therefore, the participant was left to provide an initial planning so that the automated system could resolve the remaining conflicts.

Since a human operator would make the initial planning, the algorithm did not change the arrival times of aircraft that had already received an RTA change from the operator. The operator was able to release these aircraft for automated planning. Using this technique, the operator could request aircraft to arrive earlier, generating room for the automated system to resolve conflicts behind that aircraft.

The display was almost identical to that of the initial experiment, except for:

- The addition of colours to aircraft symbols to indicate their planning state. This enabled participants to recognise which aircraft were unplanned, planned by the operator, planned by the automated system, or probed for a changed arrival time but not yet committed.

- The delay indicator would change to red when one of the aircraft in a conflict could not be delayed sufficiently to resolve the conflict by delay only. This addition aimed to counter the perceived bias toward delaying aircraft.

- Beside buttons on the screen, keyboard commands were provided for cancelling or committing probed RTAs and for releasing aircraft to the automated planning system.

4.8.2 Traffic

Only a single traffic sample was used to avoid the confound generated by differences in the traffic sample. While participants were shown the same scenario twice in that experiment, none of the participants actually noticed the similarity.

The traffic sample was generated using the same technique as in the prior experiment (see Section 4.6.1). To reduce task demand, the parameters were altered by trial and error to create a more manageable sample. While the same spacing interval of 200 seconds was used, the additional buffer was increased from 80 to 100 seconds. The standard deviation of the initial schedule error was decreased from 150 to 140 seconds.

In the previous experiment, the standard deviation of the uncertainty varied uniformly between 50 and 200 seconds. Due to this variation, differences in uncertainty between aircraft were not very pronounced (see Section 4.7.1). The resulting smaller differences in uncertainty between flights made the amount of uncertainty irrelevant in most spacing decisions. Therefore, the second experiment used two possible values for the standard deviation of the uncertainty: Either 50 or 200 seconds.

Finally, a second scenario was generated by using the same traffic sample but a smaller required spacing of 160 seconds. This scenario had a lower task demand load but was otherwise comparable to the first scenario both in schedule and prediction error.

The reduced task demand is noticable in the magnitude and number of changes to resolve the scenario: The optimal solution to the planning problem required a total of 14 minutes of RTA changes in the scenario with 160-second spacing and 24 minutes in the scenario with 200-second spacing.

4.8.3 Participants

Eight participants were randomly selected from the participant group from the previous experiment. By doing so, the participants had some initial experience with the display, and no confounding factor would be introduced by a difference in experience with the display. However, some participants indicated they had no real recollection of the previous experiment and the potential strategies.

4.8.4 Procedure

The training consisted of the same scenarios as the previous experiment, extended with scenarios that explained the automated spacing capability. Participants were coached in using the keyboard to allow focus on the time line. No coaching was performed during the measurement runs.

In the previous experiment, participants suggested that an explanation of strategies during training would have helped in understanding the provided information. Using a checklist, an instructor provided a fixed set of recommendations to each participant during the training:

- A red delay indicator means that a solution has to be found by moving aircraft forward.
- Use the automated spacing algorithm whenever possible to save time.
- The delay indicator shows which aircraft form a single conflict.
- If presented with a high amount of excess, make an initial spacing to simplify the spacing problem.
- Try to “break up” the delay indicator as this reduces a larger problem with many aircraft to several smaller problems with a few aircraft.
- Note that delaying the last aircraft under the delay indicator does not break up that indicator until it is moved completely to the end.
- It is best to reduce excess deviation such that there is a little occupation expectation above the limit. Do not try to remove all excess as prediction error is likely to either undo your correction or achieve the spacing by itself.
- When uncertainty is visualised, note that moving uncertain aircraft helps less in reducing the excess occupancy expectation than moving aircraft with low uncertainty.

4.8.5 Independent variables

Since the setup was similar, most experiment variables were identical. As the displays and scenarios had been changed, the experiment had the following independent variables:

- Scenario: high or low task demand load.
- Display: with or without uncertainty
- Participant: the individual participant in each experiment.

In all experiments, subjects would follow a fixed repetition of the scenarios. The first two runs for each participant were either with or without the uncertainty visualisation. Finally, the combination of scenario sequence and display sequence were balanced over the different participants to prevent confounding effects of training or fatigue.

4.8.6 Hypotheses

Rather than setting broad hypotheses based on the effects on the total schedule, the hypotheses for the second experiment focus on possible changes in planning strategy due to the presence of uncertainty visualisation.

The occupation expectation indicates whether a set of aircraft fits, regardless of their exact scheduling. An initial spacing can be achieved by roughly spacing aircraft such that the expectation is near the capacity. Subsequently, fine-tuning is possible near the limit for the automatic spacing algorithm.

The occupation expectation based on uncertainty indicates the risk that the initial solution will be incorrect due to prediction error. Visualising uncertainty may, therefore, help in deciding to postpone a decision until more certainty on the schedule is available.

> $H1$: Visualisation of uncertainty reduces the number of actions applied to individual flights.

Given sufficient space and information on the uncertainty, the amount of uncertainty may inform the size of the required buffer. A suitable strategy would be to increase the spacing buffer for aircraft with higher uncertainty.

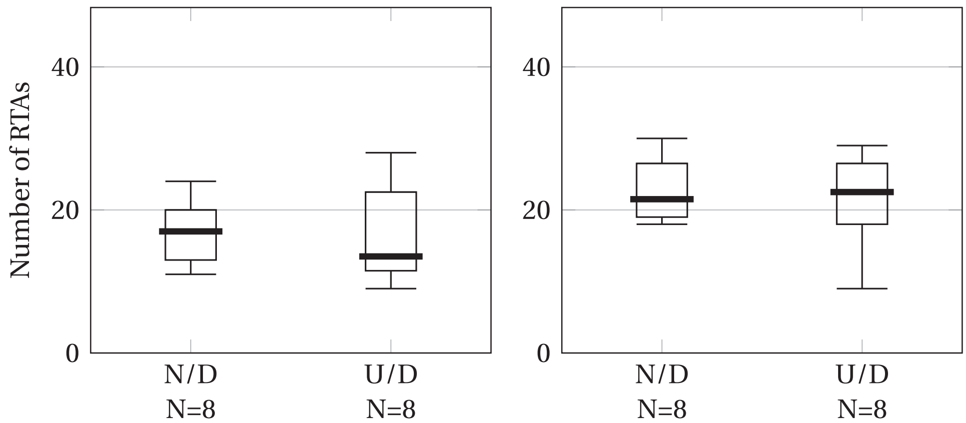
> $H2$: Visualisation of uncertainty increases the buffer applied for flights with higher uncertainty.

4.9 Second experiment results

All of the experiment's 32 runs were successfully completed. This section will first describe the objective measurements, followed by the workload analysis. The last section will draw conclusions from these results and the hypotheses.

4.9.1 Objective metrics

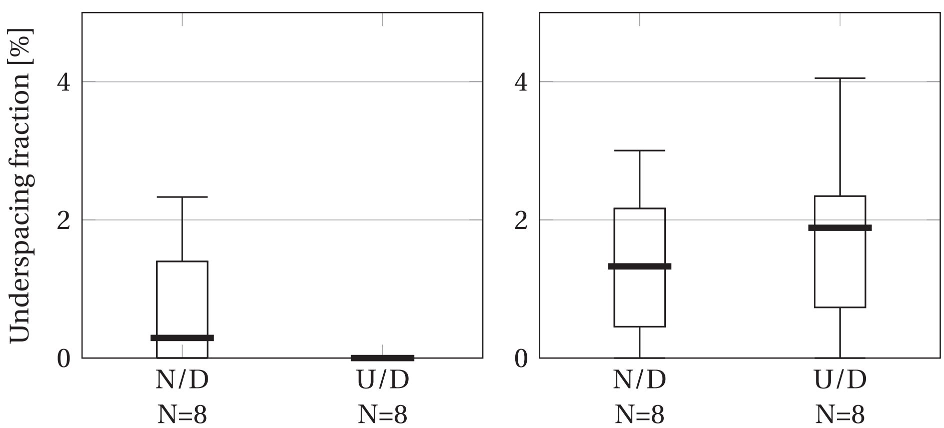
Compared to the previous experiment (Figure 4.14), Figure 4.21 shows a lower number of RTAs. The scenarios in the second experiment were longer, and therefore more aircraft were handled. Fewer small corrections were needed on the left-hand side of the time line due to the automation, therefore reducing the total.



(a) 160s spacing (b) 200s spacing

Figure 4.21: Number of RTAs requested in the second experiment.

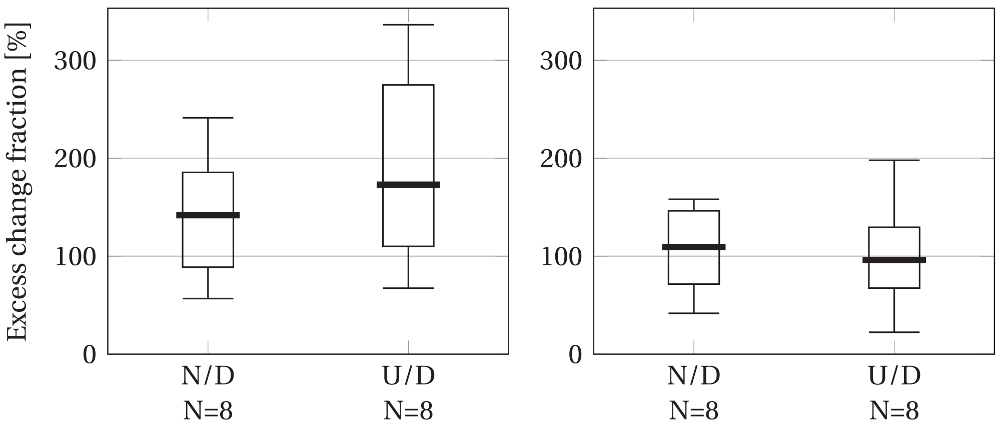
In the scenario with a lower spacing requirement—and therefore lower demand—the number of actions tends to be reduced when the display shows uncertainty although the variation tends to increase. The difference is not significant ($T = 12, p = 0.37, r = 0.08$). In the high demand scenario, the difference is less pronounced and also not significant ($T = 17, p = 0.83, r = 0.05$). While this finding invalidates the first hypothesis, the results should be evaluated in conjunction with the quality of the spacing result.



(a) 160s spacing (b) 200s spacing

Figure 4.22: Total insufficient spacing at landing. Note that no spacing errors were made in the 160s spacing scenario with uncertainty display.

The low demand scenario—in combination with the automated spacing algorithm—allowed some participants to prevent any underspacing. In combination with the uncertainty display, all participants were able to provide sufficient spacing. Even in the scenario with high traffic density, several participants delivered sufficient spacing. Compared to the low-demand scenario, however, slightly more errors were made in general, and even more with the uncertainty indicator. In both scenarios, these effects were not significant ($T = 6, p = 0.35, r = 0.24/T = 10, p = 0.50, r = 0.17$ respectively).



(a) 160s spacing (b) 200s spacing

Figure 4.23: Amount of requested deviation more than required for optimal solution.

As in first experiment, participants in this experiment requested about twice the amount of deviation than required for the optimal solution. The corrections made by the automated algorithm reduced efficiency even further but are excluded from the analysis. The uncertainty visualisation moderately increases the amount of excess change in the lower demand scenario to ($T = 2, p = 0.03, r = 0.56$), and appears to decrease changes in the high demand scenario. The latter finding is not significant, however ($T = 16, p = 0.78, r = 0.07$).

The automated algorithm performs part of the spacing task. The number of additional corrections by the algorithm, therefore, provides an extra measure of quality. Figure 4.24 shows that introduction of the uncertainty display did not affect the number of RTAs requested by the automation ($T = 11, p = 0.16, r = 0.25/T = 17, p = 0.44, r = 0.04$).

Figure 4.25 shows the spread in timing and magnitude of the RTA changes. Similar to the previous experiment, the display without uncertainty has a higher peak at the first point where the participants could request an RTA. At higher demand, more corrections occurred on the left side of the time line, with the uncertainty display leading to even more corrections.

Figure 4.26 confirms that participants tended to act later when uncertainty was provided, in particular in the high-demand scenario. The effect is not present in the low density scenario ($U = 7713, p = 0.023, r = 0.14$). In the higher demand scenario, a medium effect size is measured ($U = 10721, p < 0.01, r = 0.38$).

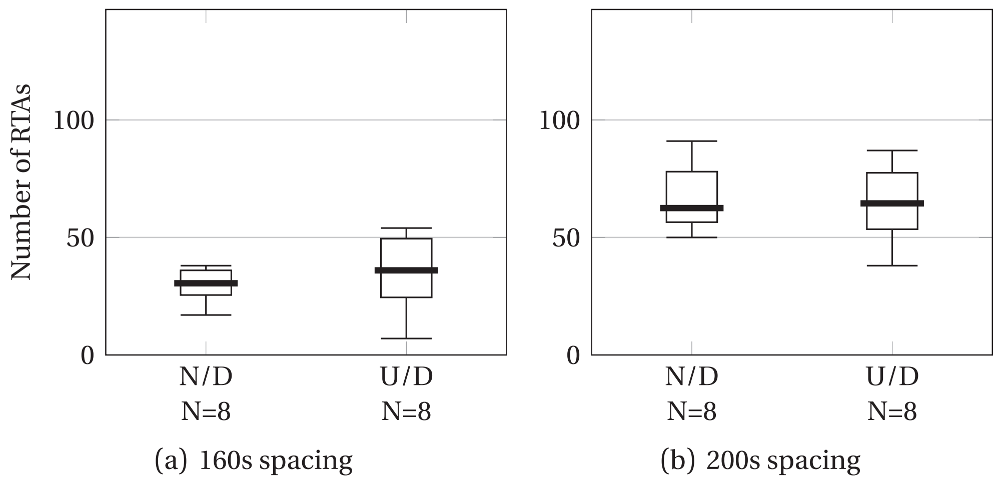


Figure 4.24: Number of RTAs by the automated algorithm.

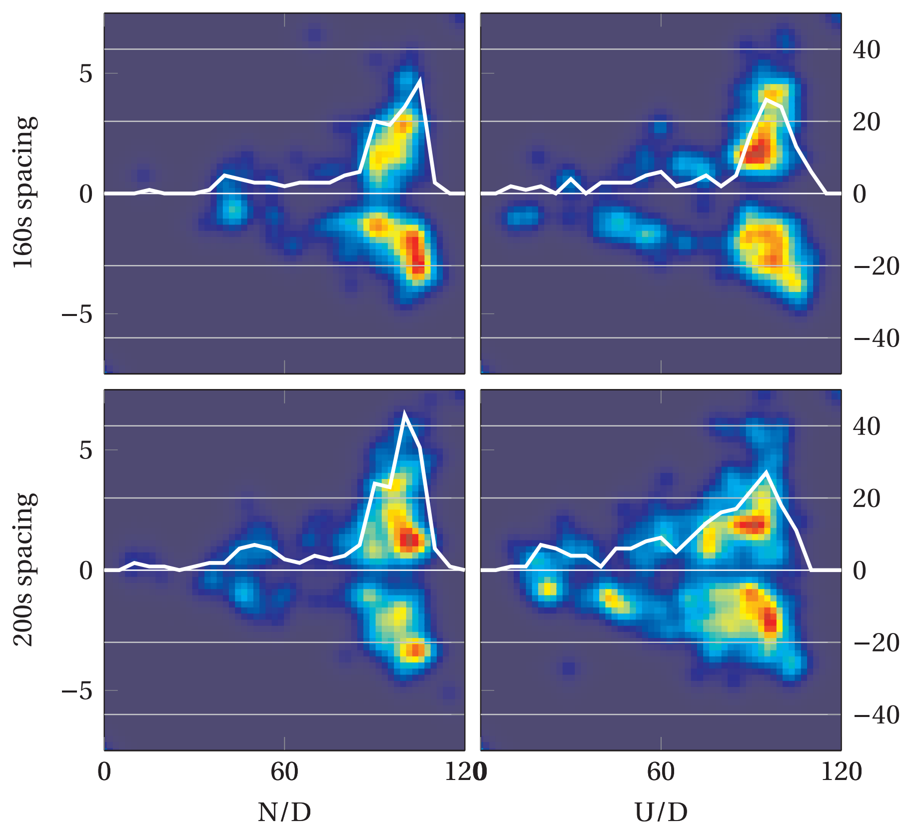
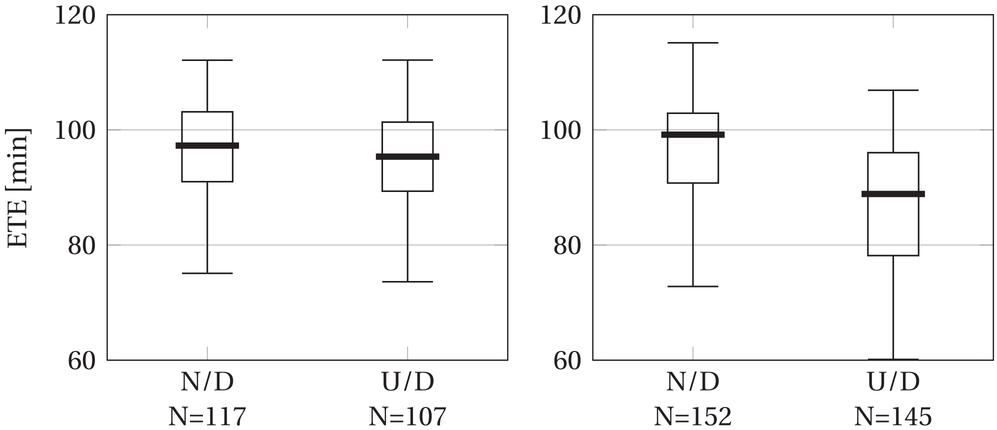


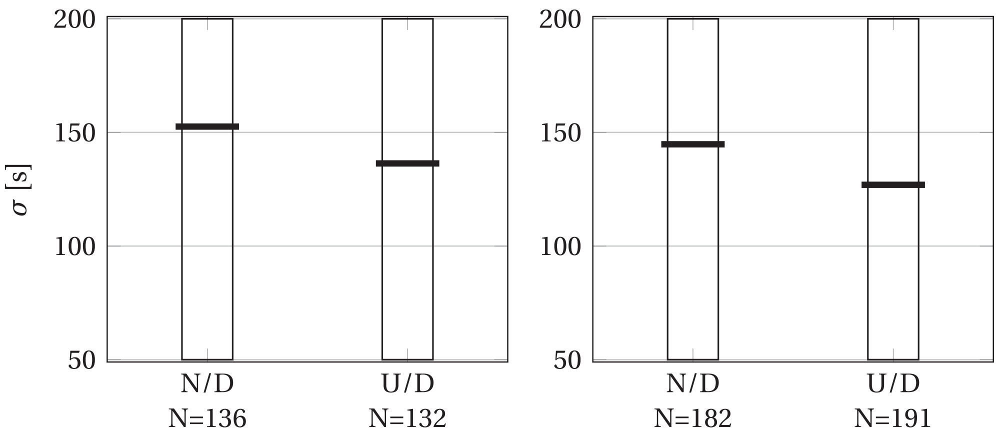
Figure 4.25: Location and ETE of RTAs. The heatmap indicates the ETE (x-axis, [min]) and amount of requested change (left y-axis [min]). The plot indicates the number of RTAs per 5 minutes (right y-axis).



(a) 160s spacing (b) 200s spacing

Figure 4.26: ETE at which RTA changes were requested.

Figure 4.25 also shows fewer actions on the left side of the time line. The smaller distribution in Figure 4.26 confirms this finding. No corrections were needed in this scenario once the automated system gained responsibility for separation. The lack of corrections during the automated phase suggests that initial spacing was sufficient for the automated system to resolve remaining conflicts.



(a) 160s spacing (b) 200s spacing

Figure 4.27: Initial uncertainty of aircraft for which RTA was requested in the second experiment. Since aircraft only started with a 50 or 200 s standard deviation uncertainty, no boxplot can be calculated. The middle indicator, therefore, indicates a mean.

Figure 4.27 shows the relation between the initial uncertainty of aircraft and the number of RTAs requested for those aircraft. According to the graph, the participants applied more corrections to aircraft with higher initial uncertainty. The aircraft with higher uncertainty are more likely to have a higher prediction error. Therefore, it is likely that these aircraft will need further corrections as they tend to deviate more.

In general, the preference for high uncertainty aircraft decreased somewhat when uncertainty was shown. Most participants appeared to prefer adjusting the more accurate flights when they were aware of the flights' uncertainty. The difference due to visualisation of uncertainty is not significant for the low demand scenario ($\chi^2(1, N = 268) = 3.36, p = 0.07, r = 0.11$) and significant but small for the high demand scenario ($\chi^2(1, N = 336), p = 0.02, r = 0.12$).

Figure 4.28 further explores this preference for aircraft with higher or lower uncertainty. The heat map indicates when RTA changes were requested for high and low uncertainty flights. The white line indicates the running mean uncertainty over 20 minutes of the aircraft that received an RTA at that time. Especially in the first quarter of the time line (i.e., 1.5-2 hours before arrival), the preference for requesting RTAs from low uncertainty flights increases.

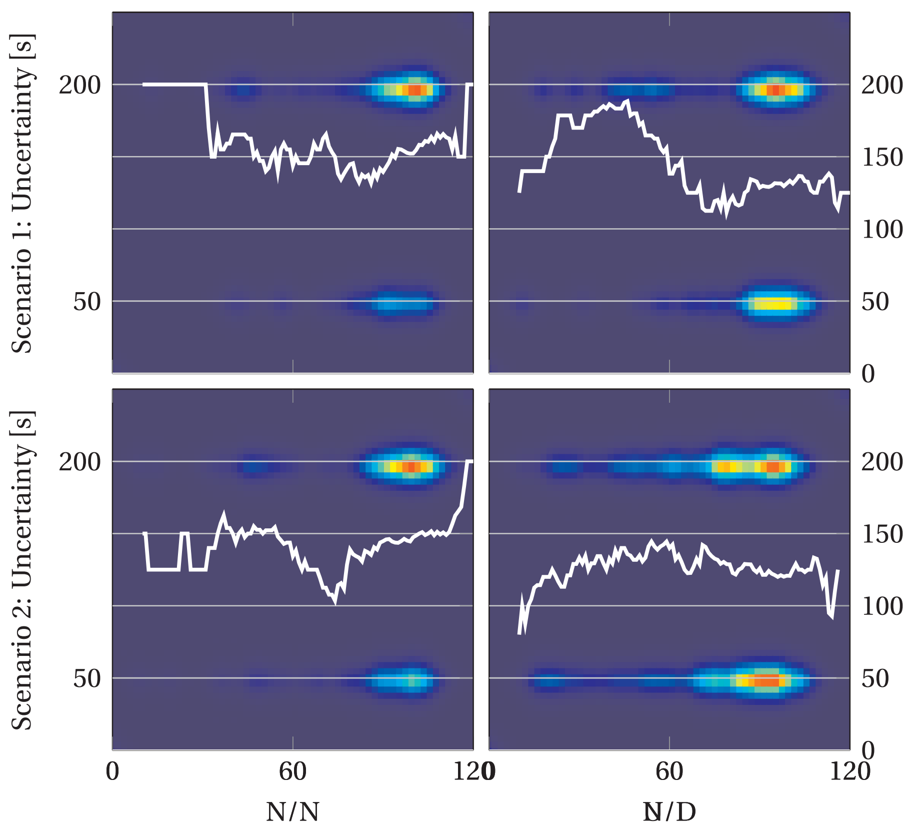


Figure 4.28: Preferences for high or low uncertainty flights in requesting RTAs. The heat map indicates at what time and at what uncertainty RTAs were requested. The white line indicates the running mean uncertainty over 20 minutes of flights for which an RTA was requested in that interval.

To test the hypothesis that visibility of uncertainty influences the buffer applied between aircraft, the remaining buffer at the 60 minutes mark is calculated. At this point, the automated algorithm has not corrected the spacing yet. Figures 4.29 and 4.30 show the distributions of the buffer with, respectively, the leader and trailer of the aircraft for which a change was requested.

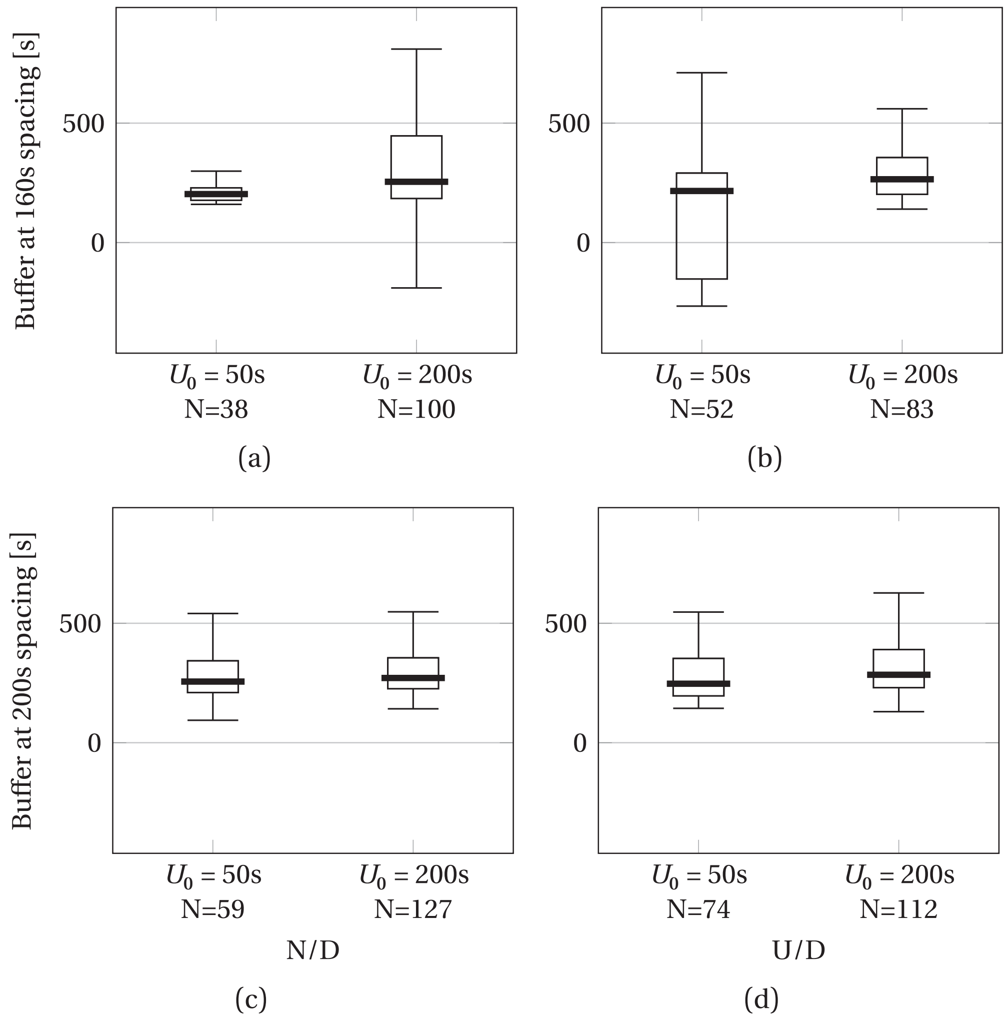


Figure 4.29: Relation between uncertainty and applied buffer with respect to leading aircraft at 60 minutes before landing. U_0 represents the initial standard deviation of uncertainty of the flight.

The difference in buffer with the leading aircraft in the low demand scenario is not significant for either the flights with low uncertainty ($U = 978, p = 0.47, r = 0.01$), or the flights with high uncertainty ($U = 4020, p_1 = 0.357, p_2 = 0.715, r = 0.03$). A similar result is found for the high demand scenario ($U = 2116, p = 0.38, r = 0.03 / U = 6686, p = 0.21, r = 0.06$). The results of the low demand scenario show a large difference in variance between the low uncertainty and high uncertainty aircraft. For the aircraft with high uncertainty, this variance is larger when the uncertainty is not shown and smaller when uncertainty is displayed.

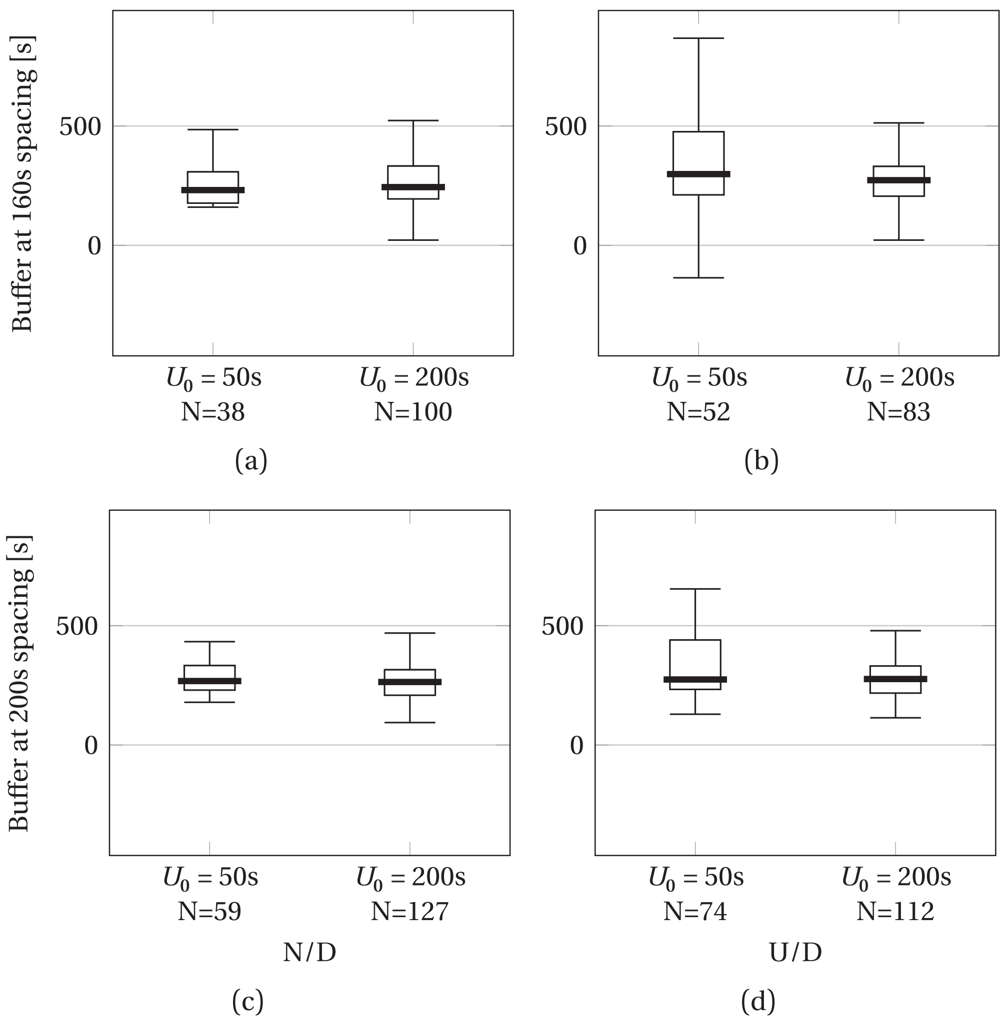


Figure 4.30: Relation between uncertainty and applied buffer with respect to trailing aircraft at 60 minutes before landing. U_0 represents the initial standard deviation of uncertainty of the flight.

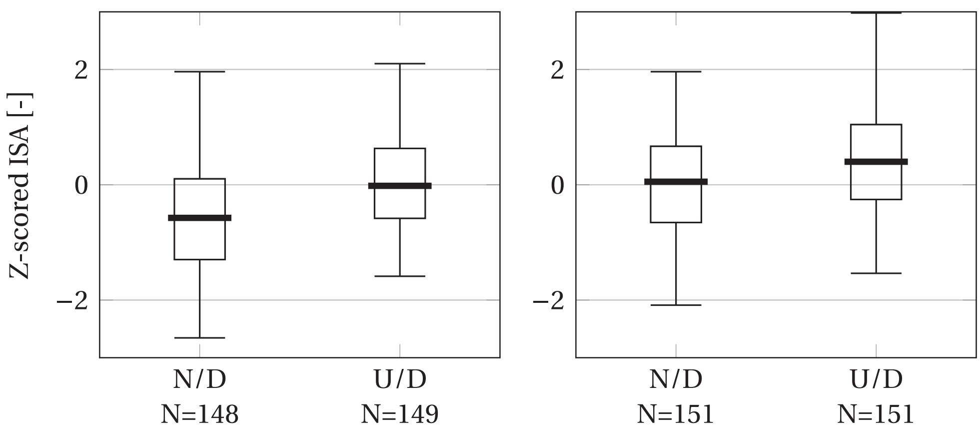
No significant difference due to the visualisation of uncertainty was found for the low uncertainty aircraft in the low demand scenario ($U = 742, p = 0.04, r = 0.25$). Similarly, no difference was found for high uncertainty aircraft in this scenario ($U = 3671, p = 0.18, r = 0.12$). The difference in variance in the spacing buffer with the leading aircraft in Figure 4.29 is not present in the spacing buffer with the trailing aircraft in Figure 4.30. For the high demand scenario, no significant effects of the visualisation of uncertainty were found either for either low or high uncertainty aircraft ($= 2017, p = 0.45, r = 0.08 / U = 6668, p = 0.41, r = 0.06$ respectively).

As shown in Figures 4.29 and 4.30, the size of the spacing buffer has no clear relation with demand, display, or the uncertainty of the aircraft under control.

However, the variation of the buffer appears to be dependent on all three, in particular for the spacing with the leading aircraft. The variation in the buffers in the high-demand scenario is smaller than the variation in the low-demand scenario. In the high-demand scenario, limited options exist to provide spacing in most cases hence not allowing for a large variety in spacing choices.

When demand is low, variation does depend on both the aircraft's uncertainty and the display of that uncertainty. At low, but visible uncertainty the strategy of the participants seems to vary strongly. Some participants provide large buffers to the low-uncertainty flights, whereas other participants provide smaller margins. Apparently, this choice also depends on the visualisation of uncertainty.

4.9.2 Workload



(a) 160s spacing (b) 200s spacing

Figure 4.31: ISA workload during the second experiment.

As with the previous experiment, the introduction of uncertainty visualisation increased the participants' perceived workload. The increase was significant but with a small effect size for both scenarios ($z = 4.92, p < 0.01, r = 0.29 / z = 4.39, p < 0.01, r = 0.25$). As expected, the lower demand scenario leads to a lower perceived workload.

4.9.3 Conclusions on the second experiment

The objective of the display is to allow users to decide between setting an RTA immediately versus waiting until more is known about the risk that prediction errors disturb a planned schedule. By visualising uncertainty, the operator was expected to better understand the risk of erroneous predictions for each particular flight. This should enable delaying decisions for flights with high uncertainty.

Participants did delay RTA decisions when uncertainty was displayed. However, the results of the second experiment do not demonstrate a clear difference in the number of RTAs when uncertainty is visualised compared to when it is not. Hypothesis $H1$ is therefore rejected.

Understanding the uncertainty in a prediction could also help in assigning spacing buffers. Providing larger buffers around uncertain ETAs could ensure that prediction error does not lead to a spacing violation at landing (Hypothesis $H2$). The experiment did not clearly demonstrate such an effect of the visualisation of uncertainty. In the low-demand scenario, the visualisation of uncertainty allowed all participants to plan a schedule that ensured sufficient separation at landing. However, this improvement is likely at the cost of more deviation than necessary for the optimal solution.

The experiment did show several changes in strategy depending on the display and the scenario demand:

1. Participants tended to wait longer before committing an RTA when uncertainty was displayed. This effect is particularly notable in the high-demand scenario.

2. When uncertainty was displayed, the participants issued more RTAs to aircraft with low uncertainty.

3. The strategy for assigning a buffer between aircraft varied between subjects, especially in the low-demand scenario. When uncertainty was not shown, the variance mainly applied to the uncertain ETAs (i.e., the aircraft with high uncertainty). With the uncertainty display, the variance shifted to the low uncertainty aircraft.

Finally, as with the previous experiment, adding the uncertainty visualisation to the display tends to increase workload. The participants indicated that recognising spacing conflicts was harder when uncertainty was displayed. Furthermore, deciding on an appropriate strategy based on the uncertainty required more effort.

4.10 Discussion

Neither experiment was able to demonstrate the expected effects of the visualisation of uncertainty or the a priori total delay. Several factors introduced too much variation for any effect to be notable. However, the experiments did demonstrate some effects of these additions. This section discusses the experiments' validity and the display's applicability in actual operations.

4.10.1 Effects on planning performance

The objective of displaying the uncertainty associated with an ETA is to enable more prudent decision-making. These changes aim to reduce the number of corrections to achieve sufficient spacing with minimal delay. The hypotheses, therefore, expected less excess deviation, less underspacing, earlier decisions and a lower number of RTA changes.

Both experiments failed to provide a definite conclusion on spacing performance. The type of visual aid did not significantly affect the amount of excess spacing or underspacing. Similarly, the number of RTA changes depended more on the scenario, not on the displays.

The participants requested RTAs later when uncertainty was provided. The participants explained that the block-wise display provided a more obvious signal for a need for deconflicting aircraft. This form of the occupancy indicators stacked more clearly than the wider occupancy expectation curves.

At high demand, the number of options to resolve spacing issues is too limited for uncertainty information to be useful in the decision-making process. Both experiments showed more substantial effects of the different visual aids at lower demand. The second experiment, in particular, showed more spacing errors and more excess adjustments with visual aids. The experiment also showed that the presence of uncertainty information changed the variation in buffer strategies between participants from the higher uncertainty flights to the lower uncertainty flights.

When decisions are based on the visualisation of uncertainty, the participants need a good understanding of the effects of uncertainty on the resulting schedule. In the first experiment, the participants explained that understanding that effect proved difficult. The changes to the second experiment did not lead to the expected improvement in performance either.

4.10.2 Expert display for expert users

Training in the second experiment did include an explanation of possible strategies. However, resulting strategies between participants did still show considerable variation (see Figures 4.22 and 4.23). It may well be that operators need a better understanding of the scheduling problem itself before information on uncertainty provides additional benefits.

Prior work on Ecological Interface Design (EID) stressed that that approach aims to develop interfaces for the control of complex systems by expert users of those systems [12], [13]. Bennet and Flach propose that, since the system is complex, the interface also has to represent that complexity. Only then does it provide all information on the controlled system that could be of value in the decision-making process [13]. In the case of our experiments, the participants may first have to grasp the intricacies of managing arrival times in uncertainty before elements on the display are meaningful to their decisions. Besides asking operational experts as participants, users may need further training to understand the unfamiliar presentation of, perhaps familiar, concepts, as suggested by Jamieson [14].

4.10.3 Validity compared to actual operation

During both the first and second experiment, the uncertainty was introduced by an error based on a fixed cumulative error based on the aircraft's uncertainty. With an uncertainty that decreases linearly with remaining time to fly, this error decreases linearly as well. In actual operation, the error can change instantly or linearly (see Chapter 3). Furthermore, the error does not necessarily change in the same direction. The application of the actual uncertainty as modelled in Chapter 3 will be addressed in Chapter 6.

The gradual deviation of an aircraft from its indicated ETA meant that participants could space aircraft and then see how the situation would develop. At no point would the sequence change suddenly. In reality, many changes will be stepwise (see Section 3.4.4). A few participants did recognise the error's linear behaviour and actively corrected for it (i.e., providing a buffer in the direction opposite to the error). A future experiment should apply a more realistic error model with more instantaneous and non-linear changes. Such a model could force participants to consider that an aircraft can change to anywhere within its predicted uncertainty envelope.

The application of accelerated-time simulation to the AMAN obfuscates one of the major problems of such a long-term planning process: Such substantial delays between decision and actual result make it harder to keep track of the original plan and to notice deviations from that plan. An experiment would see a real-time test with such delays for a valid comparison. However, this would require multi-day experiments with similarly long training sessions.

4.10.4 Extending the scope

Looking back at the abstraction hierarchy in Figure 4.9, some elements were addressed in these experiment. However, the coverage is far from complete. This section will discuss four limitations in the current display: The possible effects of the sequence on capacity, the applicability to an airport with multiple runways, the actual possibility and cost of an adjustment of arrival time, and the effects on the upstream controller's workload.

The current concept does not yet address sequencing. Depending on the airport, sequence optimisation can be a valuable means to increase capacity by using the variation in spacing time between different aircraft types (see Section 2.2.2). However, determining a suitable spacing distance for presentation is not trivial: When the uncertainty on the ETA is large enough to span multiple aircraft, the sequence itself is uncertain. When the sequence is unclear, the pairs of in-trail aircraft cannot be determined. Without a determined pair, the spacing interval itself cannot be fixed. Therefore, the resulting demand graph will have to take uncertainty in the required separation into account.

The lower right of the abstraction hierarchy in Figure 4.9 shows several aspects of the system which influence the available capacity. Especially at airports where runway combinations are varied to accommodate varying demand—such as Amsterdam—timely decisions will enable early trajectory decisions for arriving aircraft and early departure decisions for departing aircraft that compete for the use of the same runway. A key problem in providing meaningful representation of multiple resources (i.e., runways) is that a distinction needs to be made between global excesses in demand (i.e., more aircraft than the airport can handle) versus local excesses (i.e., too little spacing at one runway). The first can be directly shown on the time line by raising the capacity limit to the limit of the complete airport. Visualising demand and capacity per specific runway requires knowledge of which aircraft uses which runway.

The current approach assumes that aircraft can and prefer to make an equal adjustment in faster and slower speed. In reality, however, the available speed change may not be that flexible. Furthermore, adjustments may be available through other means (e.g., take-off delay, route adjustment). This would make the available speed envelope much more dependent on individual aircraft. Also, the display does not yet indicate the cost of a change, and, thus, the efficiency of the manoeuvre. A five-minute delay can be achieved with a minimal change in speed—and efficiency—if performed two hours before landing. A 30-minute change, however, requires considerable path changes, which may cost considerably more fuel for the same gain in time. Similarly, the cost of a delay depends on the commercial operation of the airline and is unlikely to be equal for each operator or even for each flight.

This chapter did not evaluate the spatial relation between inbound aircraft. The routes of the inbound aircraft may impact the possibilities for achieving a certain sequence. Furthermore, the complexity of achieving that sequence—and thus task demand for the upstream controllers—also drives the quality of an AMAN sequence. Inclusion of those parameters requires extension of the HMI with a spatial relation which—most likely—requires a plan view of the traffic situation as also found in Section 2.5.5.

4.11 Conclusion

The current AMAN displays support planning an arrival sequence as long as the underlying ETAs are sufficiently accurate. Errors in predicted arrival times result in unreliable Scheduled Time of Arrivals (STAs). Corrections are needed when flights deviate from the planned arrival time, which increases workload and reduces flight efficiency. The current AMAN approach is, therefore, to limit the planning horizon to that at which uncertainty is sufficiently small.

The proposed display aims to show the uncertainty to the operator such that it may form an integral part of the operator's decision-making process. By showing the uncertainty as its PDF and the resulting expectation, the display aims to provide both the uncertainty and the effect of that uncertainty on the higher level objectives. This will facilitate extending the AMAN horizon.

Initial experiments demonstrate that such additional information is only of value when the schedule offers sufficient room for spacing based on uncertainty to take place. The hypothesised improvements in efficiency were not seen, however. This is mainly due to the complexity of the display, the underlying planning problem, and the difficulty of testing novel interfaces experimentally.

4.12 Recommendations

Further experiments will be necessary to investigate the potential effects of visualising prediction uncertainty. These experiments should, first of all, allow for more training per participant to allow a more in-depth understanding of the dynamics of an arrival plan based on erroneous prediction information. Secondly, the nature of the error should be more realistic such that a plan may suddenly become unsuitable rather than slowly degrade. This behaviour would put a higher significance on the risks associated with planning uncertain flights. Finally, experiments should be performed in real-time simulations to introduce the effects of long delays between action and effect.

The next steps in developing the interface itself should focus on improving support. The current display only provides information on the quality of the schedule in terms of spacing. The display does not indicate the costs of changes and, therefore, the actual quality of the solution.

Finally, the display should be developed to widen its applicability of the display to more situations. These include support for wider variation in available solution space, multiple runways and varying spacing intervals.

4.13 Bibliography

[1] M. Tielrooij, C. Borst, D. Nieuwenhuisen, and M. Mulder, “Supporting arrival management decisions by visualising uncertainty,” in *Proceedings of the SESAR Innovation Days*, D. Schaefer, Ed., Stockholm, Sweden, 2013.

[2] A. Barff, B. Favennec, P. Conroy, *et al.*, “SESAR P05.06.04 - D28 - Preliminary OSED ed. 00.01.01,” SESAR Consortium, Tech. Rep., 2012.

[3] N. Hasevoets and P. Conroy, “Arrival Manager - implementation guidelines and lessons learned,” EUROCONTROL, Brussels, Belgium, Tech. Rep., 2010.

[4] C. Maria Andrea Almansa, A. Ross Groskreutz, T. Pütz, *et al.*, “Update of 5.6.4 OSED - Step 1,” SESAR JU, Brussels, Belgium, Tech. Rep., 2015.

[5] J. Bronsvoort, G. McDonald, M. Paglione, C. Garcia-Avello, I. Bayraktutar, and C. M. Young, “Impact of missing longitudinal aircraft intent on descent trajectory prediction,” in *30th IEEE/AIAA Digital Avionics Systems Conference*, Seattle, WA, USA, Oct. 2011.

[6] “SESAR concept of operations at a glance ed 02.00.00,” SESAR JU, Tech. Rep., 2011.

[7] JPDO, *NextGen integrated work plan: A functional outline*, 2008. [Online]. Available: `http://www.jpdo.gov` (visited on 03/01/2013).

[8] N. Hasevoets and P. Conroy, “AMAN status review 2010,” EUROCONTROL, Brussels, Belgium, Tech. Rep., 2010.

[9] J. Rasmussen, “A taxonomy for analysis of cognitive work,” in *IEEE Conference on Human Factors and Power Plants*, Monterey, CA, USA: IEEE, Jun. 1992.

[10] “D2 - The performance target,” SESAR Consortium, Brussels, Belgium, Tech. Rep., 2006.

[11] A. J. Tattersall and P. S. Foord, “An experimental evaluation of instantaneous self-assessment as a measure of workload.,” *Ergonomics*, vol. 39, no. 5, pp. 740–8, May 1996.

[12] K. J. Vicente and J. Rasmussen, “Ecological interface design: theoretical foundations,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 22, no. 4, pp. 589–606, 1992.

[13] K. B. Bennett and J. M. Flach, *Display and interface design: Subtle science, exact art*. Boca Raton, FL, USA: Taylor and Francis, 2011.

[14] G. A. Jamieson, “Bridging the gap between cognitive work analysis and ecological interface design,” in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 47, Los Angeles, CA, USA, Oct. 2003, pp. 273–277.

Chapter 5

Visualisation for multiple runways

The proposed visualisation of arrival times in Chapter 4 supports operations for a single metering point (e.g., one runway). Furthermore, although the display visualises demand and capacity, it does not indicate the efficiency of the selected planning. This chapter proposes an extension of the display that provides support for multiple runways, shows the balance between demand and capacity, and indicates the quality of the actual planning. The basis of this chapter is an existing publication. This may cause some repetition of previous chapters.
The study in this chapter was performed in tandem with the experiments described in the previous chapter. Therefore, parts of the chapters may overlap. Similarly to the last chapter, this work developed in parallel with the work in Chapter 3. The experiment in this chapter uses a simplified uncertainty model that has no relation to the model proposed in Chapter 3. This allowed for a more controlled experimental setting and limited dependency on progress in both streams.

This chapter is based on the following publication:

Paper Title	Supporting runway planning by visualizing capacity balances of arriving aircraft streams
Authors	J. de Wit, M. Tielrooij, C. Borst, M.M. van Paassen and, M. Mulder
Published in	IEEE Conference on Systems, Man, and Cybernetics (pp. 2989-2994), October, 2014, San Diego CA, USA

.

5.1 Introduction

After a small decrease in 2013 due to the economic crisis, air traffic volumes in the EU steadily increased during 2014-2019 [1]. As a result, especially larger airports in the EU have to manage a large increase in arriving and departing traffic while avoiding congestion at the airport and in its surrounding airspace. A similar situation existed in the 1990s during the development of the first operational requirements for an Arrival Manager (AMAN) [2]. The AMAN concept aims to support Air Traffic Control (ATC) in sequencing and scheduling a stream of arriving aircraft into the airport.

Some Air Navigation Service Providers (ANSPs) regard AMANs as essential tools to support the Air Traffic Controller (ATCO). In these regions, AMANs are mainly used to regulate the traffic flow into the Terminal Maneuvering Area (TMA) surrounding busy airports and to balance the inbound flow of aircraft and the available landing capacity. In other regions, AMAN systems are only used as a traffic awareness tool and provide no interaction with the inbound flow. Finally, at many, mainly smaller, airports, AMANs are not used at all [3].

Although suited for the task, the currently available AMAN systems are fairly basic: all display aircraft labels at their respective Estimated Times of Arrival (ETAs) on a time line. Chapter 4 of this thesis proposed a new AMAN interface which presents arrival time uncertainty in addition to more conventional parameters. This uncertainty information might further support ATCOs in planning decisions since it allows them to anticipate potential errors in the ETAs of aircraft.

However, when conflicting interests arise between aircraft using the available runway capacity (i.e., arriving and departing traffic), ATC will need to decide on the active runway configuration. The process of runway planning concerns deciding which runway configuration to use by balancing the *available* capacity and the *required* capacity (i.e., demand). In this way, runway planning contributes to the efficiency of the arrival stream by ensuring that all arriving aircraft can land on time. Furthermore, efficient runway planning also contributes to efficient use of runway capacity by transferring available capacity to departing traffic whenever possible[1].

This chapter presents the state-of-the-art of AMAN systems and identifies the absent support for runway planning. Based on these findings, it proposes an extension to the display presented in Chapter 4 to support the runway planning decision by explicitly visualising the use of runway capacity and the potential imbalance between available capacity and demand. Two metrics are developed that provide insight into the balance between available capacity and demand, which is the primary driver in the runway planning decision-making process.

[1] Runway planning is not to be confused with runway assignment. Runway assignment is the process of assigning aircraft to runways. Supporting this task is considered a topic for future research.

This project focuses on the case of traffic that arrives at Amsterdam Airport Schiphol. At this large airport in The Netherlands, five runways are available, and the traffic demand consists of several distinct peaks of arriving and departing traffic. Furthermore, legislative restrictions prescribe and limit the use of the available runways [4]. This combination results in several changes from a configuration of two landing runways and one departing runway to a configuration of two departing runways and one landing runway [5]. Therefore, Air Traffic Control the Netherlands (LVNL) faces a runway planning decision multiple times daily. Although the study focuses on Amsterdam, several other airports also have multiple runways and varying traffic densities over the day. Therefore, those airports likely face similar runway planning decisions.

The chapter is structured as follows. Section 5.2 briefly repeats relevant the state-of-the-art of AMAN systems from Chapter 2, discusses the arrival management process, describes the uncertainty-based AMAN, and analyses how these support the runway planning process. Section 5.3 describes metrics to qualify the balance between adjusting arrival times and adjusting runway capacity. Based on these metrics, the section then proposes the extended interface of the proposed runway planning support tool. The proposed interface has been tested in a human-in-the-loop experiment described in Sections 5.4. The results of this experiment are presented and discussed in Sections 5.5 and 5.6, respectively.

5.2 State-of-the-art

In 2010, EUROCONTROL conducted a comprehensive review of the implementation status of AMAN systems in Europe [2]. This review lists the AMAN products on the market or in development at that time. The review furthermore provides an overview of the dispersion of the various products across Europe. Most of these AMAN systems are still in use today, and Table 5.1 summarises the capabilities of, respectively, Maestro, OSYRIS, 4D-Planner, IBP, OPTAMOS, the CTAS - Traffic Management Advisor (CTAS-TMA), and the newly developed AMAN proposed in Chapter 4, concerning runway planning.

These capabilities all involve determining and visualising the assigned runway, the Estimated Time of Arrival (ETA) and the Scheduled Time of Arrival (STA) of aircraft at the runway, the Time to Loose (TTL) advisories given to aircraft when a delay is required to assure proper spacing of aircraft, and the Time to Gain (TTG) advisories when aircraft need to speed up [2], [6]–[16]. The remainder of this section presents a more detailed overview of how the state-of-the-art AMANs perform on the various aspects of runway planning. It then discusses some rules of thumb used for runway planning, revealing how ATCOs tend to use their AMAN in the current operation.

Table 5.1: Capabilities of state-of-the-art AMAN systems with respect to visualising Estimated Time of Arrival (ETA) and Scheduled Time of Arrival (STA), Time to Loose (TTL) and Time to Gain (TTG), and assigned runway (RWY).

System	**ETA/STA**	**TTL/TTG**	**RWY**
MAESTRO	-	✓	✓
OSYRIS	✓	✓	✓
4D-Planner	-	✓	✓
IBP	✓	✓	✓
OPTAMOS	-	-	✓
CTAS-TMA	✓	✓	✓
New AMAN	✓	✓	-

5.2.1 Visualising arrival times

Most of the operational AMANs (e.g., Maestro, OSYRIS, 4D-Planner, OPTAMOS, and CTAS-TMA) show one or more vertical time lines on which aircraft labels display the flight number, and in some cases, the required TTL or TTG [2]. The bottom of each time line represents the current time. The position of each label along the time line indicates the ETA or STA of the corresponding aircraft. The current AMAN at LVNL, the Inbound Planner (IBP), also uses such a vertical time line, but primarily lists a textual representation of all aircraft, including flight number, assigned runway, and ETA at the runway threshold [2].

ATCOs can perform the spacing task by inspecting the space between aircraft labels along the time line using the current time line displays. However, the traditional systems do not explicitly visualise the spacing in terms of the use of available capacity. Instead, the use of capacity and associated potential delays only become evident after—in most cases automated—initial scheduling. Such an initial planning is only possible when the predicted arrival times are sufficiently accurate. If the runway decision is required at a longer time horizon, an initial schedule will be needed at that horizon. If the ETAs have a too high uncertainty, initial scheduling is not possible. The controller then has to perform the cognitively challenging task of deciding whether the available runway capacity is sufficient to accommodate all aircraft without exactly knowing when these aircraft arrive.

The new AMAN (Figure 5.1) described in Chapter 4 provides such an a priori view on how the current spacing of aircraft affects the available capacity. It shows the expected runway occupancy and the potential delay necessary to resolve the shortage of capacity, while not requiring the aircraft to be planned or to be predicted accurately.

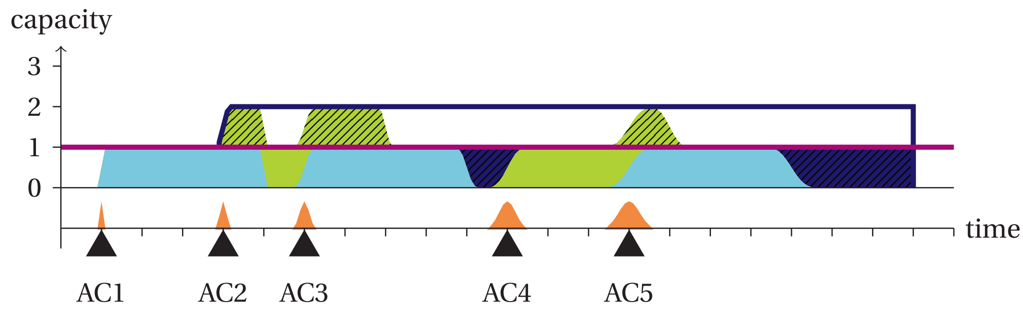


Figure 5.1: Presentation of expected excess demand, and the time required to resolve these problems by delaying aircraft as explained in Chapter 4.

The blue areas show the expected occupancy of available landing capacity. These areas show the expectation that arriving aircraft use the available landing resources at any given time. The expected occupancy of each aircraft is cumulated to indicate the expected need for capacity at any given time. When this expectation exceeds the available capacity (i.e., the green dashed areas above the red capacity line), an excess in expected occupancy is present in the arrival stream. Assuming delay as the only strategy, the excess expected occupancy can only be resolved when an equal amount of expected free runway has passed. This moment can be determined in time, resulting in an expected delay for the involved aircraft.

Chapter 4 described the initial tests of the new display in an accelerated-time experiment with student participants. This experiment did not yet show a benefit of the presentation of uncertainty. Based on input from the participants, the complexity of the display, combined with limited training, may have obscured any effect of the displays themselves. However, the experiment suggested that visualising a priori cumulative delay helped subjects provide sufficient separation without increasing the number of corrective actions.

The uncertainty presentation indicates the need for a runway configuration change to resolve capacity shortages. However, it does not yet support deciding between a runway configuration change and adjusting the arrival time. When aircraft are inbound to an airport, they are assumed to have a certain optimal arrival time: Delay is mostly regarded as a negative factor, whereas flying faster to arrive earlier leads to a higher fuel burn. When assuming the ETA to be the optimal arrival time, most AMANs visualise the deviation of this time for individual aircraft through the TTL and TTG advisories. These are the consequence of automatically applied scheduling. However, none of the AMAN systems present an overall value for the deviations from the original ETA, making it hard to judge the overall efficiency in the arrival stream. It could well be that all aircraft are spaced appropriately, but a different sequence or spacing would result in a better operational performance.

5.2.2 Runway planning support

The proposed AMAN only visualises traffic to one distinct runway. Hence no runway planning support is available at all. Other AMANs considered in this research indicate on which runway arriving aircraft are scheduled to land, as assigned by approach control, based on the active runway configuration and the direction from which the aircraft arrive. Some systems use different time lines for different runways (e.g., Maestro, OSYRIS, 4D-Planner, OPTAMOS), whereas other systems use different areas in the interface to list traffic for different runways (e.g., IBP).

As discussed previously, both arriving and departing traffic may require runway capacity at the airport. Assuming single modes of operation, ATC should minimise the use of runways for arrivals to prevent delays on departures. Hence, to achieve an optimal arrival planning that considers the different traffic flows, usage of runway capacity is an important parameter. Although runway assignment is visualised in the various interfaces, the already applied automatic scheduling and assignment may hide how the current runway assignment affects the available runway capacity. Consequently, any inefficient use of runway capacity is difficult to detect. Mental processing of the information is required for this purpose, which generally takes longer and could increase the ATCO workload.

5.2.3 Runway planning rules of thumb

Two interviews were conducted with ATCOs responsible for arrival management at LVNL. A live demonstration of IBP, the AMAN in use at LVNL, and a structured interview afterwards revealed that creating a good arrival planning is a cognitively challenging task. Moreover, it became apparent that these controllers use several rules of thumb to supplement the information shown on their AMAN. For example, specifically for runway planning, one such rule states that a TTL advisory of four minutes or more warrants using another runway configuration to increase available capacity. Hence, it is the ATCO who integrates the information into the decision to change the configuration, not the AMAN. Suppose the display would show the ATCO the available capacity and the demand. In that case, a more accurate decision could be made and a runway planning closer to the optimum may be achieved.

5.2.4 Concluding remarks

Most systems present the ETA, the required TTL/TTG, and the assigned runway. Those systems support the runway assignment process by showing the assigned runway for each aircraft. However, none of the current systems indicate the total deviation from the optimal arrival times and, therefore, the solution's efficiency. Furthermore, the systems do not explicitly visualise the efficiency of using the runway and require the ATCO to infer the balance of capacity and demand from the TTL/TTG. This balance is the critical control parameter that governs the runway planning problem.

5.3 Runway planning support tool

Before assigning flights to a runway, ATC will have to decide on the runway configuration to be used. State-of-the-art AMAN systems do not directly support the runway planning decision. Although most systems support visualising traffic assigned to multiple runways, none explicitly supports the controller in deciding whether and when to change the active runway configuration. Hence, a mental process is required to integrate the information into a decision, which could increase the workload of the ATCO, and the decision taken will most likely be sub-optimal. Runway planning concerns the decision which runways to use for arriving traffic. The new AMAN should at least support the use of multiple runways and support the decision when to use these runways.

5.3.1 Performance metrics

When multiple runways are available for landing, the runway planning decision (i.e., when to open or close the secondary runway) needs to be considered. Following the findings in Section 5.2, three metrics are defined to drive the runway planning decision: *Absolute Total Deviation* (ATD), *Unused Runway Time* (URT), and *Total Time Deviation* (TTD).

The ATD is defined as the sum of the absolute TTL and TTG advisories resulting from arrival management actions, as assigned to all individual arriving aircraft. This metric assumes that any deviation from an undisturbed arrival time decreases flight efficiency. In its current form, the ATD also assumes that a specific duration of delay is equally inefficient as the same duration that the aircraft arrives earlier than scheduled.

The URT is defined as the time that a runway is reserved for arrivals but not used to land aircraft. This metric is based on the assumption that every second a runway is reserved for arrivals but not used, the runway is unavailable for departing traffic and causes one second of unnecessary departure delay. For this research, it is assumed that one runway is dedicated to accepting arriving traffic, and one runway is in alternating use for either arriving or departing traffic, as is the case at Amsterdam.

The purpose of the interface is to support the controller in deciding whether to use additional runway capacity—and possibly generate unused runway time—or to let aircraft deviate from their optimal arrival times and thereby add to the total arrival time deviation. When the balance between capacity and demand is presented as the balance between ATD and URT, the quality of that balance can be modelled as the sum of these two (TTD). An optimal solution would then minimise the sum. This approach assumes that one second of idle runway is equally undesirable as one second deviation from the scheduled arrival time.

5.3.2 Interface description

Figure 5.2 shows the interface of the new AMAN augmented with runway planning cues. It is an extension to the AMAN interface for a single runway described in Chapter 4. The projected arrival schedule is shown on a horizontal axis representing time. Aircraft symbols on this axis travel from the right to the left, where the right-hand side indicates the outer prediction horizon and the left-hand side represents arrival at the runway.

The new interface consists of three time lines. The lines represent the total airport capacity—top—and the time lines of two individual runways. The available capacity for both runways is one unit, as indicated by the red horizontal line. A possible shortage of capacity on one runway might be resolved by changing the arrival times of aircraft in the arrival stream for that runway or by assigning aircraft to the other runway. The grey area in this figure represents the time that the second runway is not used for arriving aircraft. Hence it is available for other purposes such as departures.

The top time line shows the total available capacity and the total expected occupancy for all active runways (i.e., the airport). The total occupancy time line reveals whether the available capacity at the active runway configuration is sufficient or a configuration change is required. When the occupancy level exceeds the available capacity, and this problem cannot be resolved by issuing TTL or TTG advisories, the only remaining solution is to change the runway configuration.

5.3.3 Supporting runway planning

The total airport picture allows direct visualisation of the efficiency of the presented solution as a function of time. The operator can immediately detect the size and sources of inefficiencies by showing the cumulative contributions to the ATD and URT. The cumulative totals at the prediction horizon indicate the total efficiency of the current schedule.

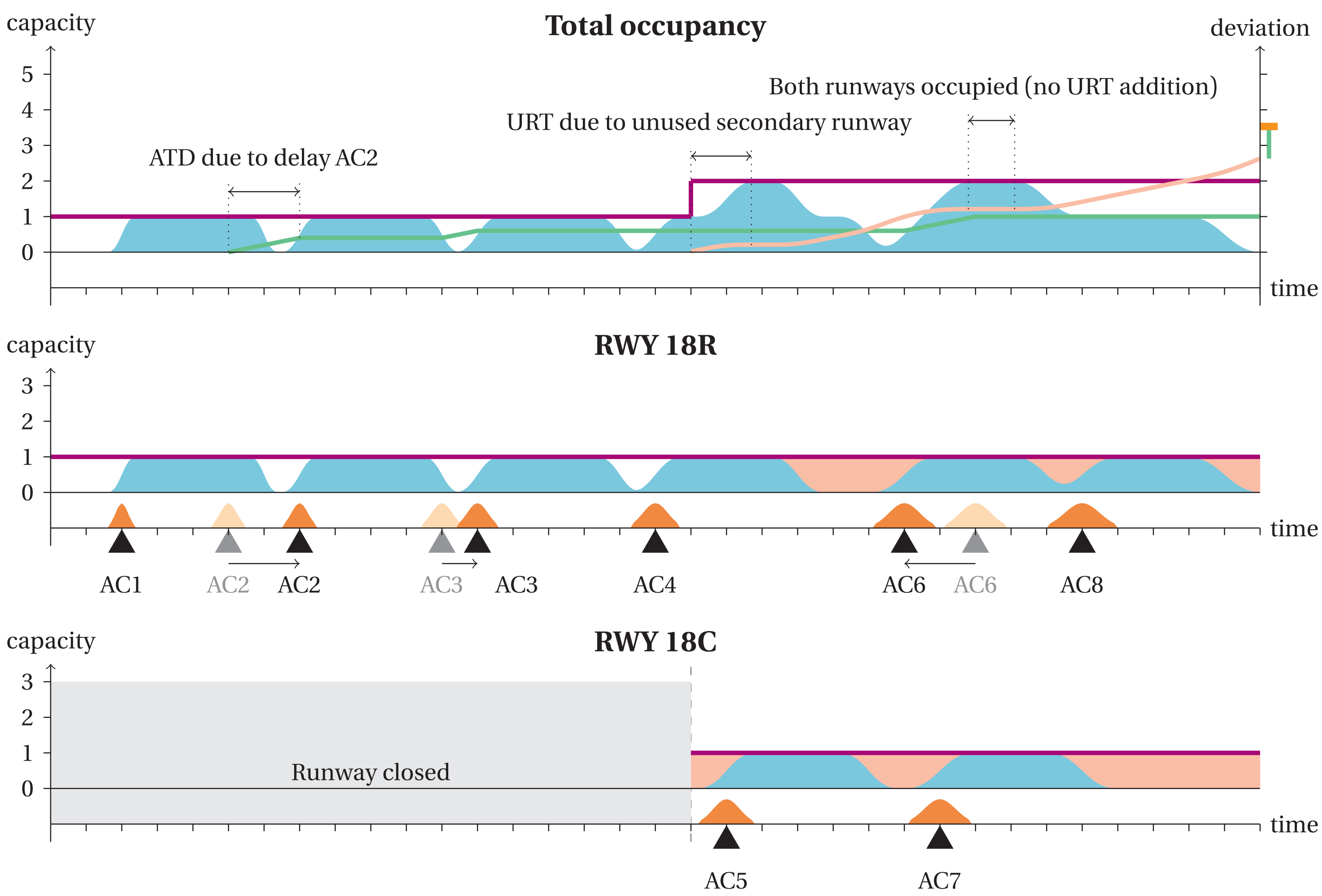


Figure 5.2: AMAN interface with runway planning support. Two runway time lines (middle/bottom) show arriving aircraft with their arrival time uncertainty and expected runway occupancy visualised. The total occupancy time line (top) shows the total arrival time deviation (green graph), unused runway time (red graph), and total performance deviation (sum of ATD and URT, orange indicator to the right of the top right vertical axis, with green line showing the magnitude of the smallest contribution to the sum).

Figure 5.2 visualises the ATD as a green line on the graph. For clarity, the figure shows the origin of each increase in deviation on the primary runway time line. On this line, some aircraft have a TTL or TTG advisory, as indicated by the arrows underneath the aircraft symbols. The rise in the green graph corresponds to the absolute amount of deviation given to the aircraft below. The rise spans the time in which the arrival time has been shifted. The rise ends at the new position of the aircraft symbol. At every point along the arrival stream, the green graph shows the total arrival time deviation of the planning up to that point.

The red graph on the total occupancy time line in Figure 5.2 provides a graphical representation of the URT. The rises in the graph correspond to the unused capacity—the area not covered by blue expected occupancy—of the runway time lines below. For clarity, these areas have been shaded in pink in Figure 5.2.

Usually, at least one runway handles arriving traffic at Amsterdam. During operations with a single landing runway, the URT would increase when no aircraft uses the runway. This increase would decrease the quality of the balance of demand and capacity while no demand is present. Therefore, URT is only counted when the controller decides to reserve additional runway capacity for arrivals. In this way, the unused runway time becomes an indicator of the efficiency of using an additional runway.

The TTD, and therefore the quality of the total balance, is indicated by an orange indicator on the right vertical axis of the total occupancy time line. The bar indicates an optimal planning when it is at the origin and lower quality as it moves higher. Since the bar indicates the sum of two graphs, the magnitude of the two components may be lost to the observer. Therefore, a coloured vertical line reveals the magnitude of the smallest of the two contributions. The line connects the graph with the largest component of ATD or URT to the marker for the TTD. The colour of the line matches that of the smaller component. When the line is green, deviations from the optimal arrival times provide the smallest contribution to the total performance deviation (i.e., the contribution of ATD is smaller in Figure 5.2). When the line is red, the more substantial part of the total deviation is caused by these deviations in arrival times. This total performance deviation indicator aims to support the ATCO in making the trade-off between deviating aircraft from their optimal arrival time or using an extra runway.

5.3.4 Expected Benefits

The AMAN with visualisation of uncertainty as proposed in Chapter 4 provides an a priori view of the potential delay in solving a predicted schedule. By combining this view with the ability to change runway capacity, the new display should allow runway configuration decisions even when the arrival times are uncertain. Using the combined information, the runway configuration decision may be taken earlier than currently feasible.

The approximate timing of the runway decision can be based on the overall expected delay available from the new AMAN. However, the exact timing of a configuration change then depends on the final schedule. The presentation of the quality of the schedule can support fine-tuning the choice between adjustment of arrival times and the time of the configuration change. By providing such support, the workload of the ATCO is expected to decrease. At the same time, the efficiency for airspace users, both arriving and departing, is expected to increase.

5.4 Experiment

The display was tested in an initial experiment with 12 novice participants. The purpose of the human-in-the-loop experiment was to test the effect of the visualisation of total arrival time deviation, total unused runway time, and total performance deviation on the quality of the runway planning decision.

5.4.1 Method

Similar to the experiments on the basic display concept described in Chapter 4, the new visualisation was tested in an accelerated-time human-in-the-loop simulation. Participants were tasked with planning traffic to a fictional airport with two available runways in the experiment. The primary objective to be achieved was to provide sufficient spacing between consecutive aircraft at arrival (i.e., the left side of the time line). The secondary objective was to minimise the total deviation (TTD).

Initially, only one runway was available. As with the basic display, the controller could assign Required Times of Arrival (RTAs) to the inbound aircraft to provide spacing. Additionally, in this experiment, the controller could open and close a secondary runway and assign aircraft to that runway.

5.4.2 Participants

As with the basic display concept (see Section 4.6.1), the display is not necessarily specific to aviation. Secondly, the simulated Air Traffic Management (ATM) concept is currently not operational, i.e., active controllers will not have any experience with such a concept. Therefore, initial validation of the display not necessarily requires operational ATCOs.

Twelve participants have taken part in the experiment. All participants were students or faculty from the faculty of Aerospace Engineering and had no experience and, at most global domain knowledge of arrival management. Five subjects had taken part in the initial experiment on the basic display [16][2].

5.4.3 Scenarios

Three scenarios were designed to provide an equal amount of necessary runway changes and comparable complexity. The scenarios were run at 30 times actual speed to provide sufficient task load and measurement data. Each scenario lasted ten minutes, representing five hours of arriving traffic in real-time.

Traffic in the simulation contained three distinct traffic peaks in which the use of a secondary runway was unavoidable. All aircraft were assigned a default runway to which the flights would be assigned when both runways were available. The distribution of the aircraft was generated using the algorithm described in Section 4.6.1 with an average spacing of 200 seconds between flights and a 20-second buffer. Deviation from the regular distribution followed from a normal distribution with a 200-second standard deviation over ten aircraft. The schedule contained a continuous flow of aircraft to the main runway. Peak demand was simulated by generating traffic on the secondary runway during the peak period using the same algorithm. During a complete run of a scenario, between 69 and 74 flights would land.

The experiment used the same model for ETA uncertainty as used in the first experiment described in Section 4.6. The model consisted of a normal distribution with a standard deviation that decreased linearly to zero as aircraft approached the airport. The initial standard deviation was randomly assigned between 50 and 200 seconds.

Prediction error was based on the uncertainty to ensure consistent behaviour of the flights. The error was selected as a random point between 0.1 and 0.9 on the cumulative density function of the aircraft's uncertainty.

5.4.4 Experiment variables

Independent variables

The first independent variable was the status of the runway planning Decision Support Tool (DST) (i.e., whether it was turned on or off). With the DST *on*, the lines showing ATD and URT and the indicator of their sum were visible.

A second independent variable was the scenario under test. The three scenarios were generated to have comparable complexity and task load. As demonstrated in previous experiments, however, the scenario may have a considerable effect on the experiment outcome (see Section 4.7.1). Table 5.2 shows the six conditions

[2] This experiment was developed and executed in tandem with the two experiments described in Chapter 4.

executed in this experiment, covering all combinations of DST status and scenario. The experiment conditions were applied using a Latin square design to balance and reduce the effects of training and fatigue.

Table 5.2: Experiment conditions based on combinations of independent variables.

Condition	DST status	Scenario
A	On	1
B	On	2
C	On	3
D	Off	1
E	Off	2
F	Off	3

Dependent measures

The primary dependent measure in this experiment is the TTD as a consequence of arrival management. Since this deviation is the sum of ATD and URT, the latter two were measured during the experiment to determine the total performance deviation. Per scenario, the minimum performance deviation has been calculated beforehand. An optimisation algorithm assessed the spacing for each aircraft and subsequently determined the schedule with the smallest TTD. This value is then used as the baseline against which the measured data is compared.

In a successful implementation, the controller's workload should not increase when the runway planning DST is active. The workload has been measured using an Instantaneous Self Assessment (ISA) [17] on a continuous visual analogue scale that popped up every thirty seconds. A sound was produced to prompt for a reply when the bar appeared, and after five seconds, the outline started blinking. Both the ISA response and the response time were measured. The analogue scale appeared above the time line and did not occlude the display or the buttons. Although participants were instructed to respond to the ISA as soon as possible, their primary attention lay with the planning task.

5.4.5 Setup and procedure

The simulation was provided on a PC with a 24-inch screen and mouse interaction only. Participants used headphones to reduce disturbance from outside noise and to provide the audible cue for the ISA measurement.

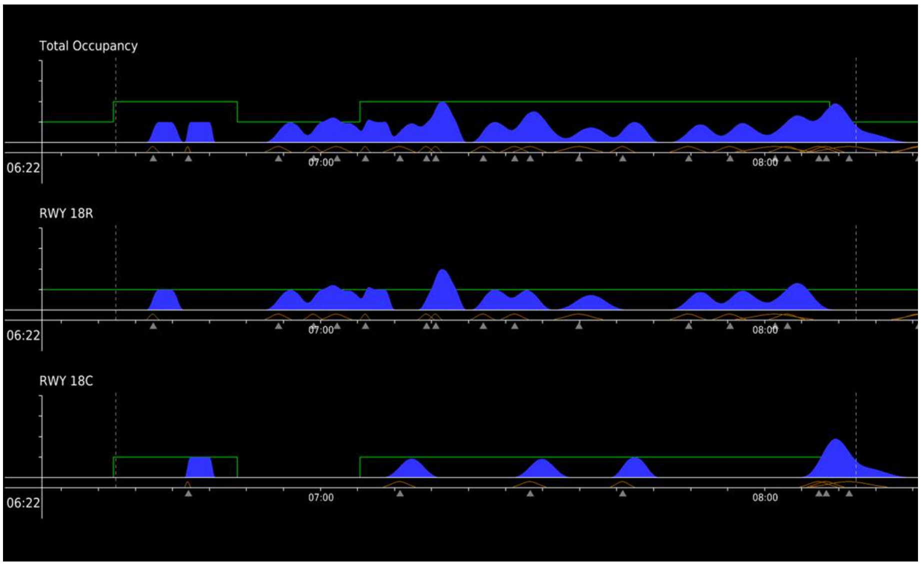


Figure 5.3: Experiment display with multiple runways.

Subjects were able to select aircraft using the mouse. On selection, the aircraft symbol, its uncertainty and the contribution to the expected use of capacity were highlighted. A different RTA could be probed by dragging symbols. Aircraft could also be assigned to the other runway by dragging the symbol to the runway. Assigning aircraft to the other runway could only be performed at times when the runway was opened.

The secondary runway was opened by clicking and dragging the capacity at the desired time. Operationally, the opening and closing of runways and the associated shifting of traffic patterns take time. A changed runways status had a minimum duration of 20 minutes to emulate this constraint.

Both for RTAs and runway status, any change was deemed provisional until explicitly confirmed by clicking a button on the screen labelled *Confirm* using the mouse. A second button—labelled *Cancel*—allowed removal of the probed changes and to the last saved state.

Earlier experiments on the new AMAN display showed considerable training effects. To reduce such effects in this experiment, participants first followed a step-by-step method explaining the various components of the interface. Subsequently, participants performed twelve complee training runs with increasing traffic complexity using one display version. After completion, the same procedure was repeated for the other display type.

5.4.6 Hypothesis

The baseline display in Chapter 4 aimed to support planning arrival times such that the spacing requirement is met with minimal deviation through RTAs. However, the display did not indicate the amount of deviation imposed on flights. Through the introduction of a second runway, an optimal balance has to be achieved between deviating aircraft from their ETAs and reserving a secondary runway without using it.

The objective of the performance deviation indicator is to support improving the total performance deviation. This deviation—ATD + URT—is measured, and the sum is expected to decrease when the information is provided. The main experimental hypothesis is, therefore:

> *Visualisation of the performance deviation indicators leads to less deviation from the optimal scenario value, thus improving the overall runway planning decision.*

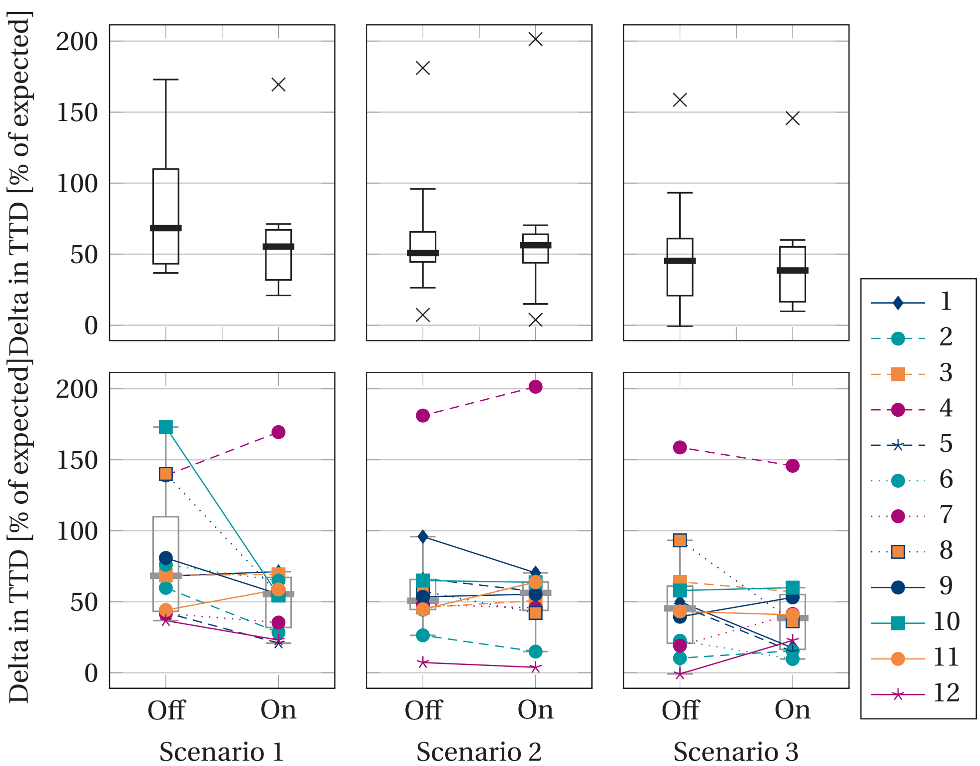


Figure 5.4: Difference of TTD against expected TTD for all 12 participants. The bottom graphs provide the same box plots with each participant's individual contributions, including the individual changes in performance deviation.
In this and all following figures, the label *Off* represents the display without the deviation indicators and *On* represents the display with those indicators.

5.5 Results

Figure 5.4 shows the distribution of the total performance (ATD + URT) when the last aircraft landed. In this and all subsequent figures, the difference performance is expressed as a percentage of the theoretical optimal performance for the scenario. The figure provides the overall distribution of the performance in the top graphs and the individual participants' performance in the bottom graphs. The bottom graphs allow visual inspection of change per participant. The colour and marking of each of the lines indicate an individual participant. A high outlier is visible in all but one condition. The plots show that those outliers all belong to a single participant.

This section further explores the effects of displaying the ATD and URT on the participants' performance. The analysis is based on a graphical presentation of the results. Any differences are analysed using a Wilcoxon signed-rank test at a significance level of 0.05 [18].

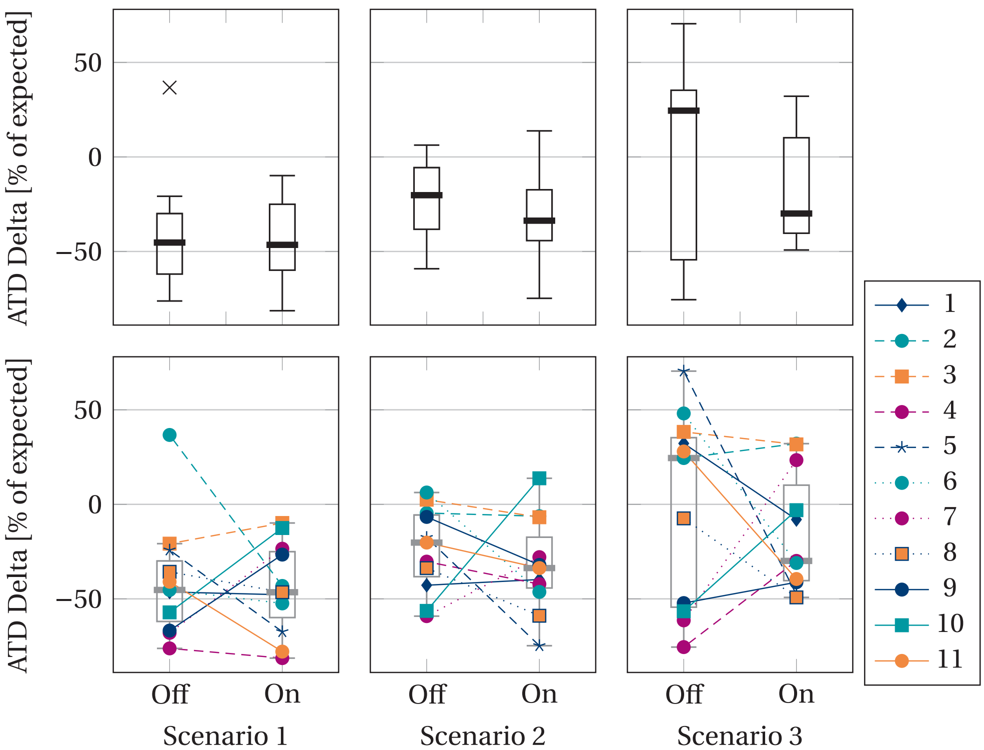


Figure 5.5: Difference with expected arrival time deviation (ATD).

5.5.1 Total deviation

Especially in Scenarios 1 and 2, most participants achieved a better total performance when using the performance indicator. This improvement is shown in

Figure 5.4 in the bottom row by the descending lines. In Scenario 3, several participants performed better while others performed worse. While total performance tends to improve with the deviation indicator, the difference is not significant ($z = 1.71, p_2 = 0.087, r = 0.20$).

TTD is the sum of ATD and URT. To further explain the different effects, Figures 5.5 and 5.6 show performance for each individual contribution. Figure 5.5 shows that most participants tended to assign far less arrival time deviation than expected in Scenarios 1 and 2. In Scenario 3, the use of the deviation indicator leads to a smaller variation. For all scenarios, and in particular, for Scenarios 2 and 3, the assigned deviation decreases. The difference is not significant ($z = 1.18, p = 0.239, r = 0.14$).

Figure 5.6 shows that the participants made more use of the additional runway than expected in the optimal solution for all scenarios and in particular for Scenario 1. The graphs also show that the outliers that are seen in Figure 5.4. As the bottom row of figures show, these are mainly due to the relatively high amount of excess use of the additional runway by one participant. In Scenarios 1 and 3, the display with deviation indicator decreased the variation in URT between participants. However, no clear effect in URT due to the presented visualisation exists ($z = 0.39, p = 0.694, r = 0.05$).

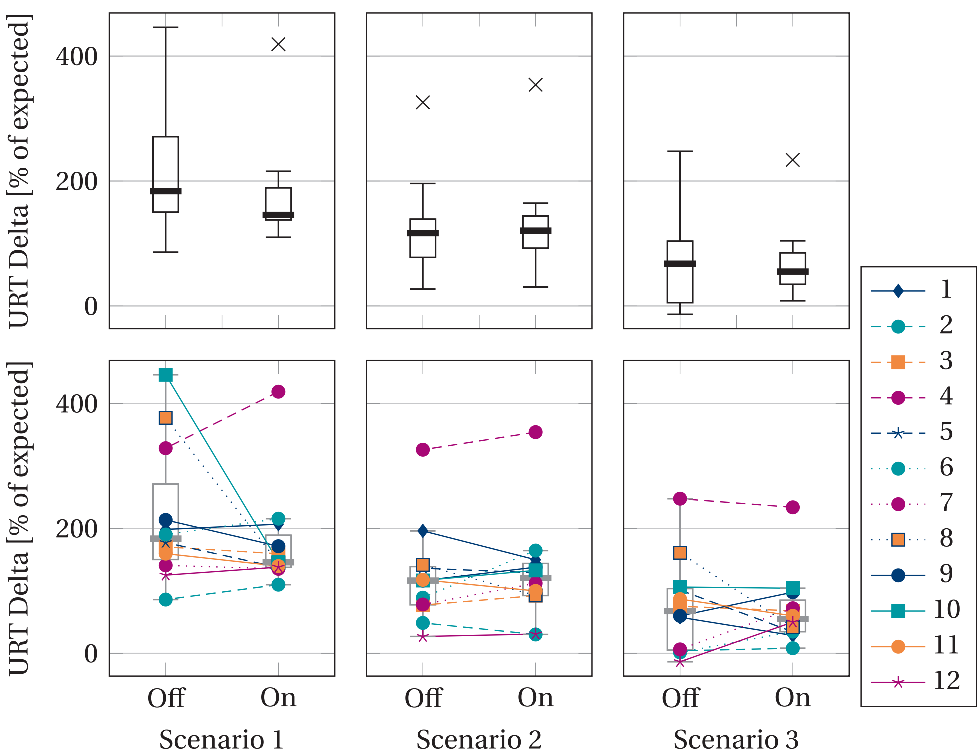


Figure 5.6: Difference with expected unused runway time (URT).

The better-than-expected performance shown in Figure 5.5 and the worse-than-expected performance in Figure 5.6 indicate that the subjects prioritised avoiding deviation (ATD) over avoiding the use of the secondary runway (URT). Furthermore, the effect of the presence of the deviation indicator appears to be dependent on the individual participant.

Differences in strategy may be based on prioritising on of the two components of the total performance.In Figures 5.7 and 5.8 a subset of participants has been selected based on the change in performance for, respectively, ATD and URT. In both figures, only those participants are included for whom performance improved for either the ATD or the URT when using the deviation indicators in at least two scenarios. Since some participants improved on *both* ATD and URT, these groups are partially overlapping.

Figure 5.7 shows the performance for the participants who showed an improvement in ATD. As expected, ATD improves while URT shows little change. In Figure 5.8, the participants with an improving URT are selected, which leads to an expected improvement in URT. For this group, the performance on the amount of deviation in ATD mostly increased.

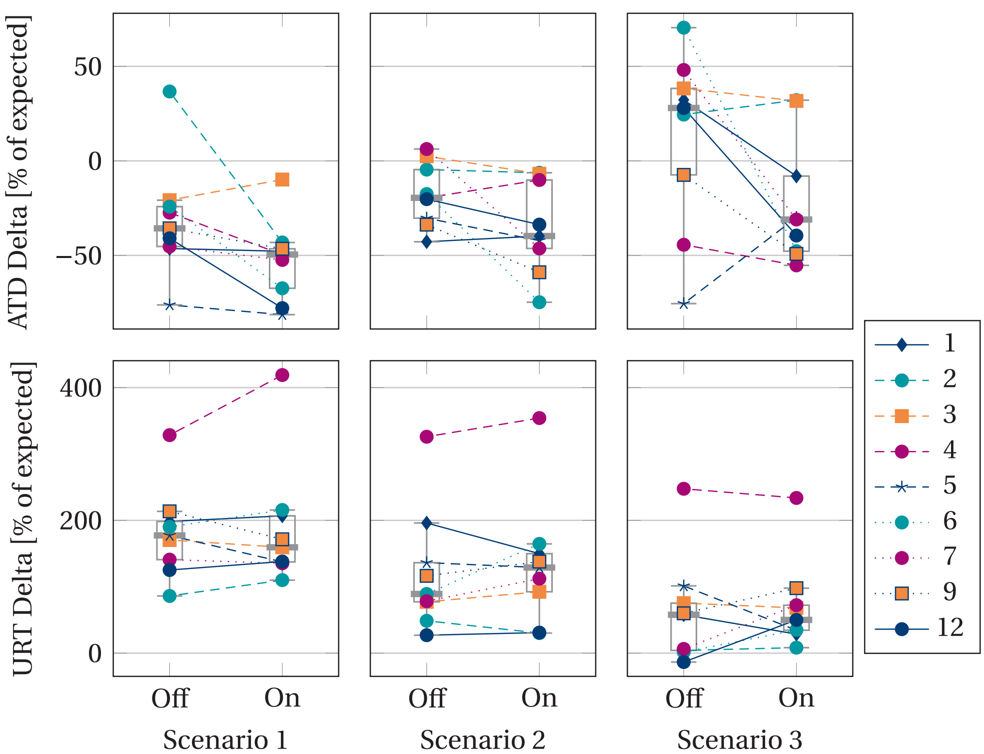


Figure 5.7: Performance changes for subjects with improving ATD in at least two of the three scenarios.

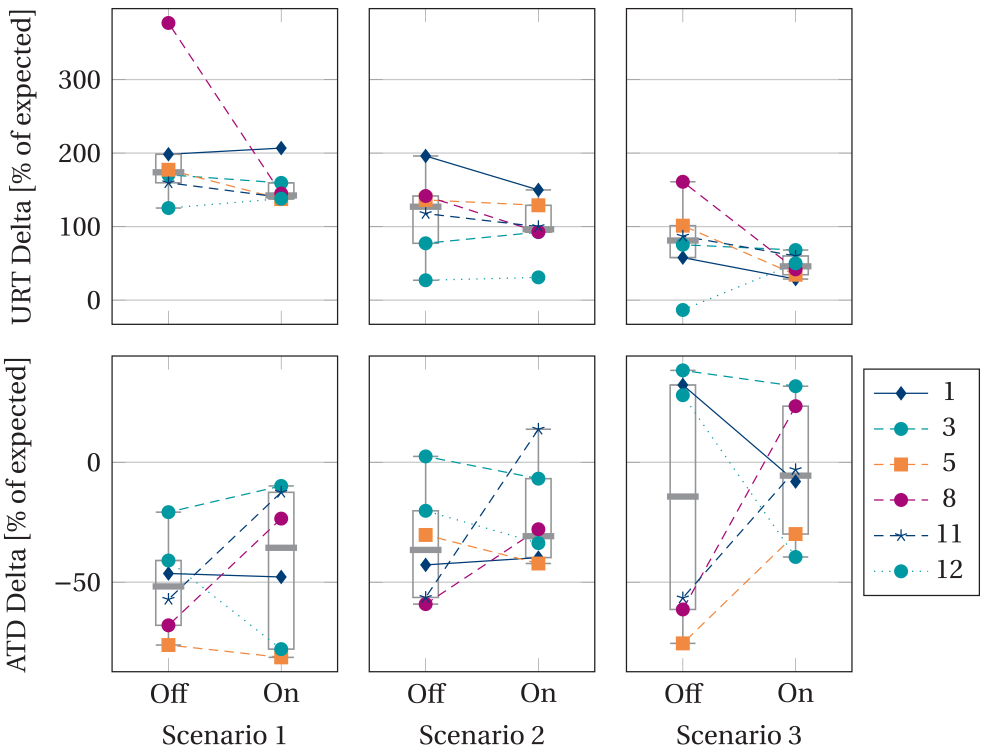


Figure 5.8: Performance changes for subjects that improved URT in at least two of the three scenarios.

5.5.2 Deviation consistency

Figures 5.9 and 5.10 present the dynamics of the ATD and URT over time in the three scenarios. These figures illustrate the increase in deviation from optimal performance for all 12 participants, as shown in the top-left graph of Figure 5.9. Effectively the heat maps show the degree of concentration of the 12 graphs of the individual participants for each scenario and each case. This concentration shows to what degree these participants performed similarly.

The heat maps support the observation that visualisation of performance deviation in the AMAN interface tends to decrease the variance of the total arrival time deviation and total unused runway time, particularly for the ATD. Throughout the scenarios, the spread of the data points increases when the performance deviation indicators are inactive (see the left- hand plots in Figures 5.9), whereas the spread remains more concentrated when the performance deviation is visualised in the interface. In other words, the presence of the performance deviation indicator tends to lead to a more homogeneous performance among participants.

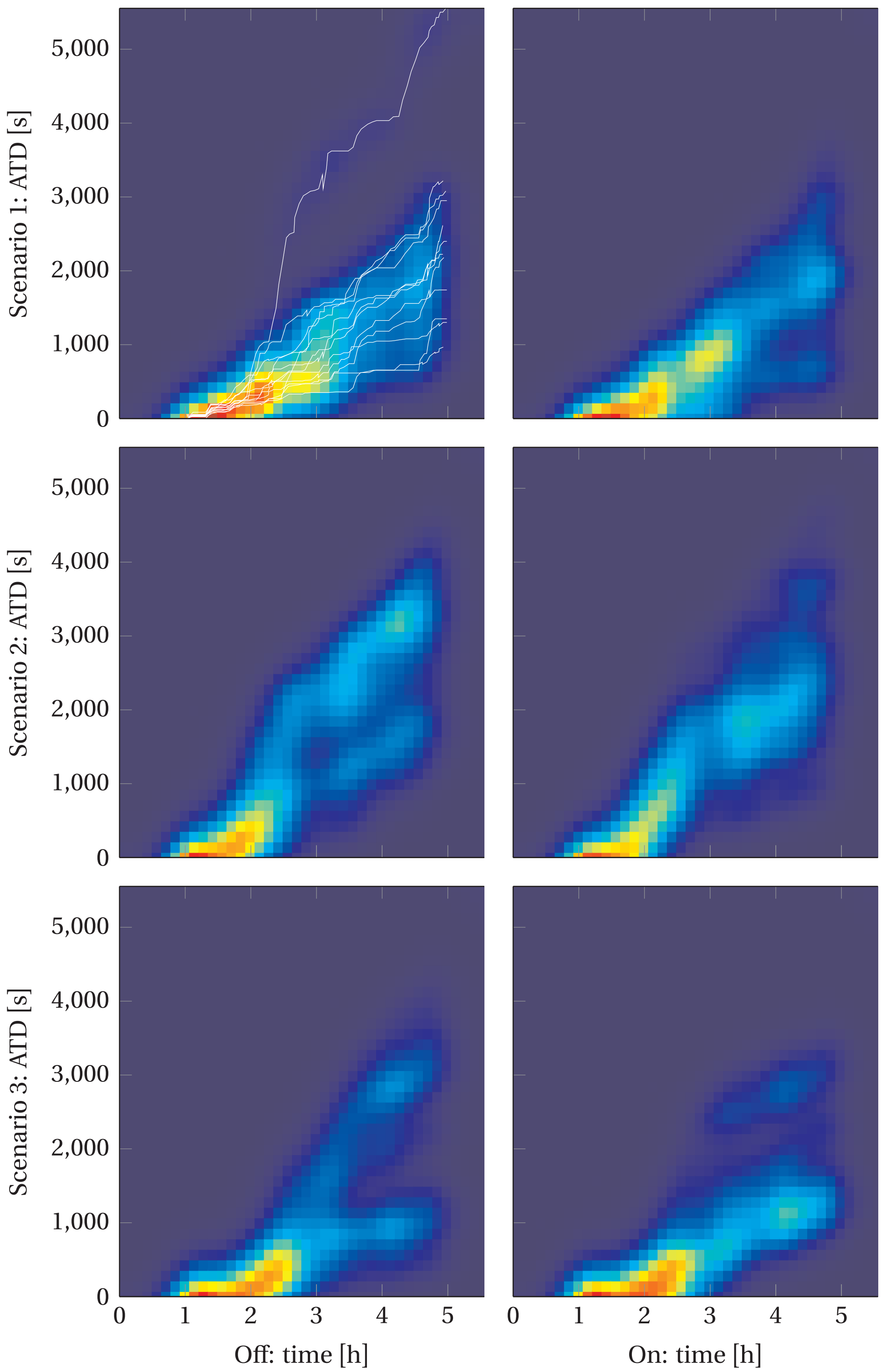


Figure 5.9: Density plot of relation between actual time of arrival and cumulative ATD at that point in the simulation.

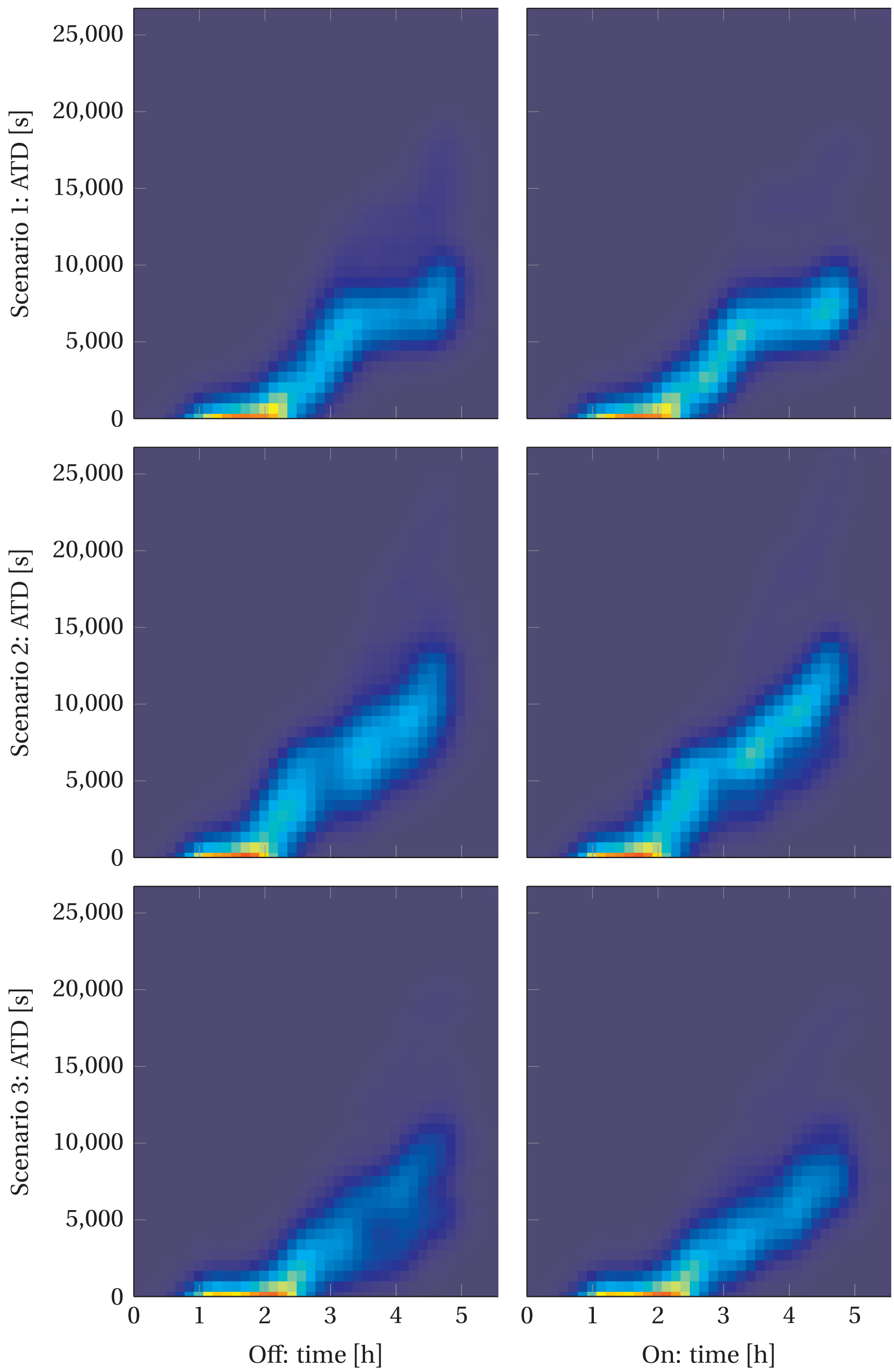


Figure 5.10: Density plot of relation between actual time of arrival and cumulative URT at that point in the simulation.

5.5.3 Spacing

The visualisation of ATD, URT and TTD are designed to improve spacing by improving the efficiency of the selected solution. The additions are not designed to reduce spacing violations. Similar to the discussion in Chapter 4, improved planning performance is only meaningful when not increasing spacing violations. Figure 5.11 provides the number of times that two aircraft landed too close after each other (i.e., the number of violations). Figure 5.12 provides the sum of the underspacing (i.e., indicating the severity of the violation).

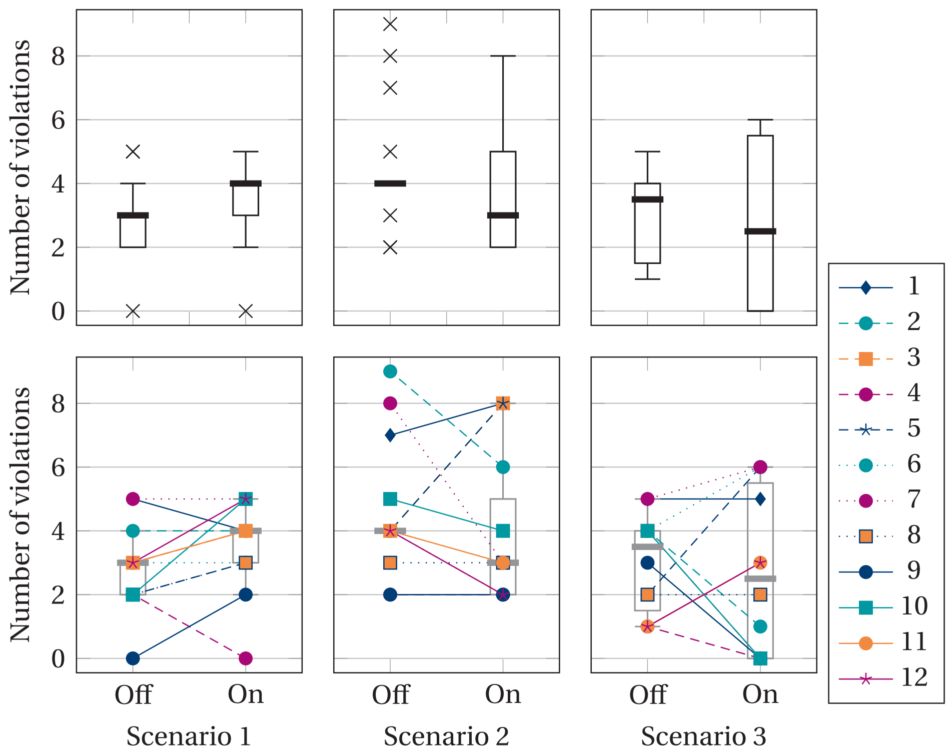


Figure 5.11: Number of spacing violations at landing.

The number of spacing violations per scenario varies between 0 and 9. In each scenario, between 69 and 74 flights landed. Of these violations, 50% were less than 12 seconds and 90% less than 60 seconds of the required 200 seconds. The display type has a small effect on the number of spacing errors, as shown in Figure 5.11. The number of violations increases somewhat in Scenario 1 when the deviation indicator is provided. In Scenarios 2 and 3, the number of violations decreases, with several participants able to avoid any spacing violation in Scenario 3. No significant differences are found, however ($z = 0.05$, $p = 0.964$, $r = 0.01$).

The total amount of spacing violations shown in Figure 5.12 is not dependent on the display. The error distribution mimics the distribution of the number of spacing errors except for two outliers in Scenario 2. Since the errors for those participants are not shown in Figure 5.11, these outliers most likely represent a single large spacing error. The differences seen in Figure 5.12 are also not significant ($z = 0.65$, $p = 0.518$, $r = 0.08$).

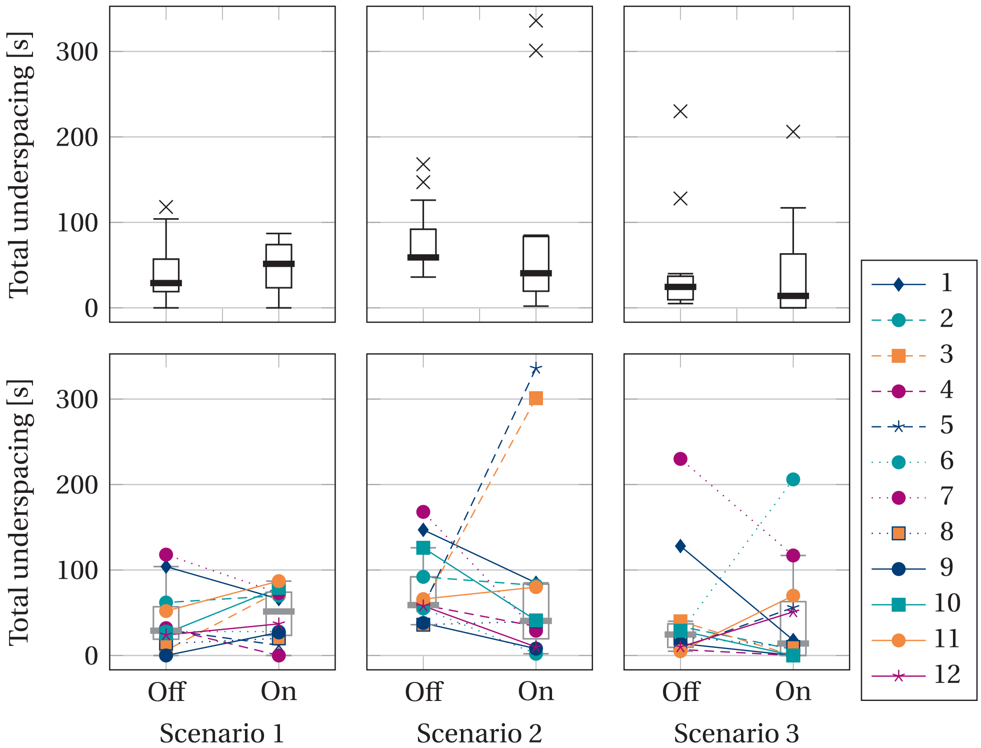


Figure 5.12: Total number of spacing violations at landing.

5.5.4 Control strategies

While not significant, Figure 5.6 indicates some changes in URT when the performance indicator is visible. Table 5.3 indicates the total number of changes to the runway configuration by all participants. The table shows how often the secondary runway was opened or closed by the participants and at what position on the timeline this action occurred. The scenario's each contained three peaks for which the secondary runway should be opened, and closed once traffic reached low levels again. The combination of three peaks in each scenario and twelve participants implies an expected total of 72 runway changes (opening and closing) per scenario.

The first observation is that the participants applied more changes than those 72 in each scenario. Furthermore, when the deviation indicator is on, more actions are performed at the left-hand side—shorter before arrival—of the time line and less at the right-hand side. Apparently, participants opted to delay their choice in runway changes when the deviation indicator was provided.

Table 5.3: Distribution of the runway changes. The total number of changes, the number of changes on the left-hand side of the time line (0-60 min) and on the right-hand side (60-120 min) are shown. The value between parentheses indicates the number of changes in which the secondary runway was closed.

Scenario	Off			On		
	Total	<60min	>=60min	Total	<60min	>=60min
Scenario 1	74 (33)	19 (9)	55 (24)	81 (45)	36 (19)	45 (26)
Scenario 2	88 (36)	24 (14)	64 (22)	97 (49)	38 (19)	59 (30)
Scenario 3	85 (43)	31 (13)	54 (30)	92 (45)	33 (13)	59 (32)

A second observation is that participants performed more runway actions closely before arrival when the deviation indicator was provided. Especially the latter observation is counter to the objective of the new display; it is intended to allow earlier changes to improve efficiency. Participants indicated that one of the objects for these actions was to improve the score when it was sure that the secondary runway was no longer needed.

Figure 5.13 further details the runway opening and closing actions with respect to the horizon on the time line. The figure splits the result between group of the participants with a strong improvement on ATD versus those with a strong improvement on URT. The figure shows that the secondary runway was closed more often when the indicator was provided. While both groups have a comparable number of runway openings, the group with an improved URT more often decided to close the secondary runway, especially on the left-hand side of the time line. The deviation indicator may have prompted the participants to close the runway when it no longer proved necessary.

Figure 5.14 provides the same graphs as Figure 5.13 but focuses at the RTAs. Here, the RTAs have been separated by their effect on the ATD: The figure distinguishes between whether the action increased or decreased the ATD. The two left columns show the change in behaviour of the participant group which showed an improvement in ATD. The number of times that these participants increased the ATD is lower when the deviation indicator is visible. The number of times that these participants reduced the ATD has not changed. These results suggest that participants who improved on ATD used the additional information to reduce the number of deviations from the optimal arrival time. The second participant group—for which the URT improved—shows no clear changes in planning behaviour regarding the ATD.

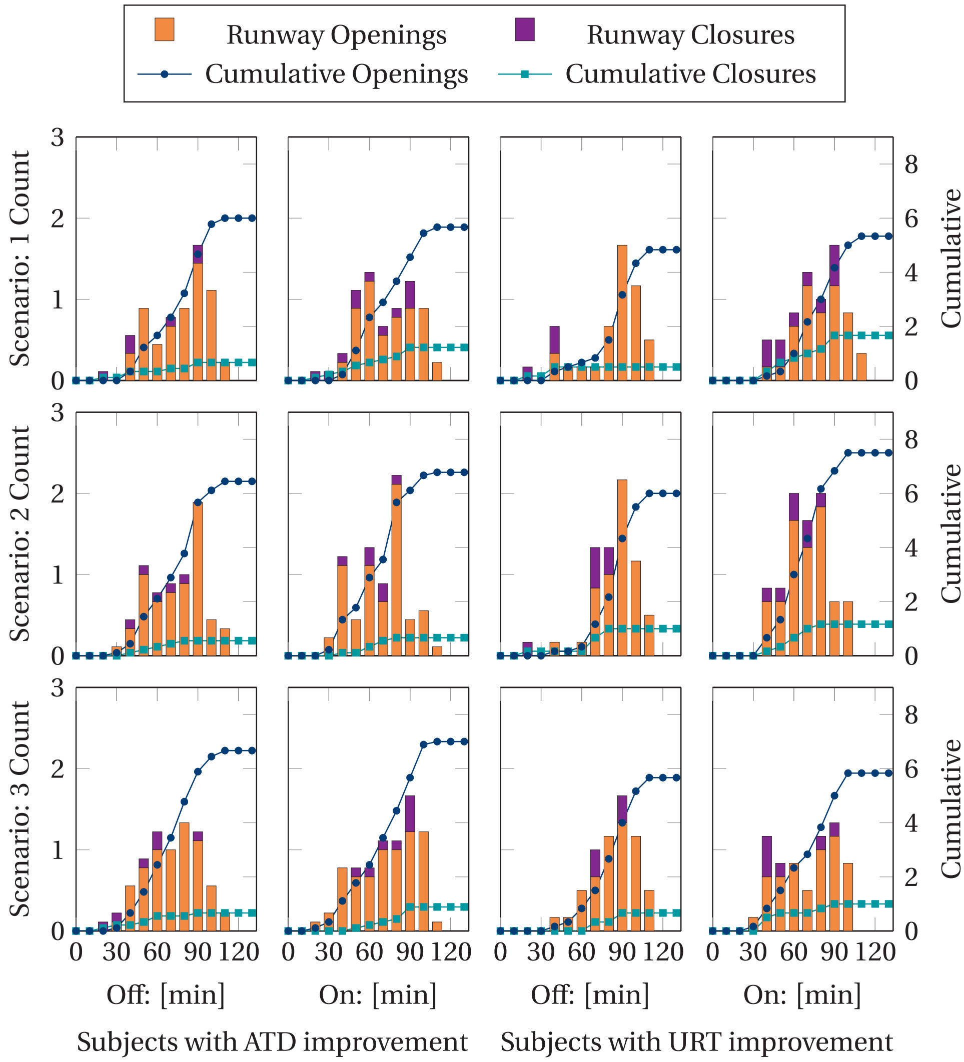


Figure 5.13: Distribution of openings and closures of the second runway on the time line. The graph shows the average number of runway changes per participant at a given point on the time line, and the cumulative total at that position

5.5.5 Workload

Figure 5.15 shows the normalised workload ratings of all participants. The boxplots show no indication of any differences in subjective workload with or without the additional display information. This suggests that the display has no effect on the workload experienced by the participants in the experiment.

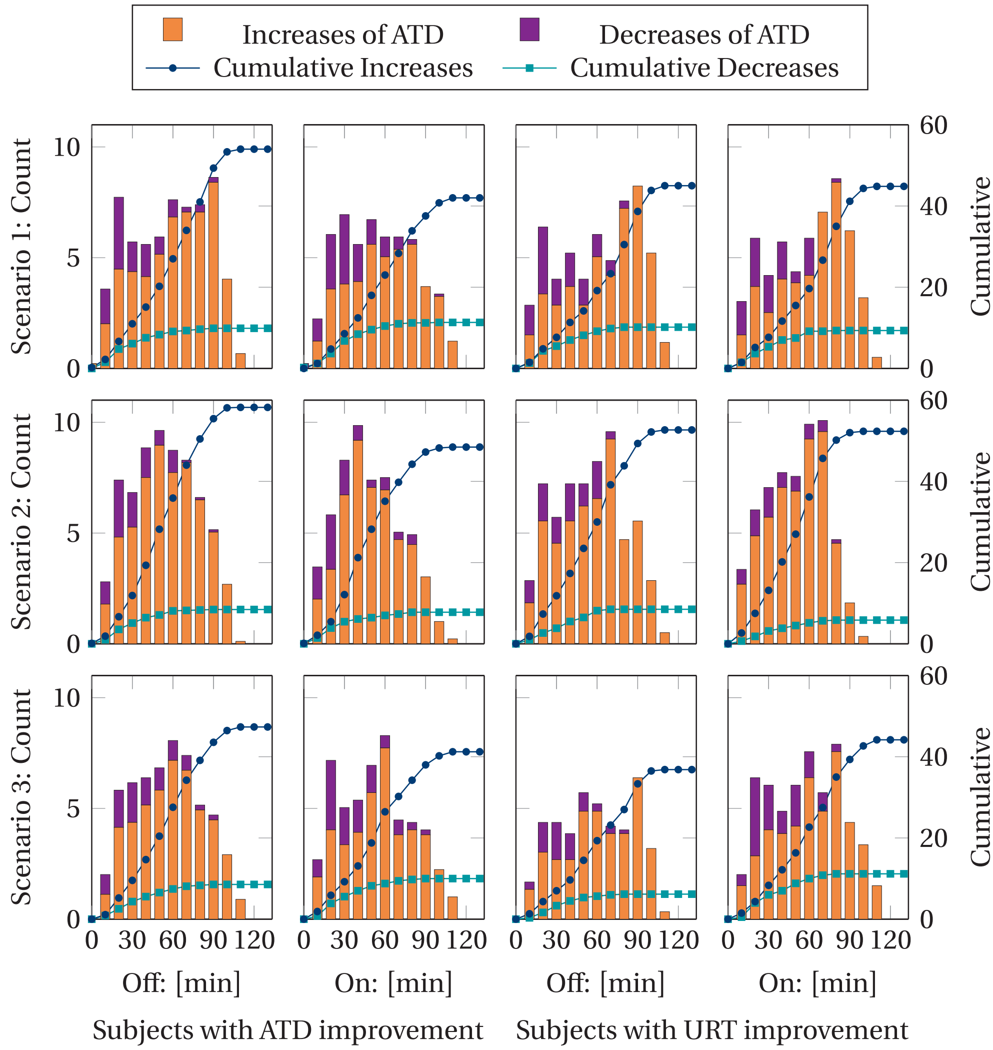


Figure 5.14: Distribution of RTAs on the time line. The graphs show the average number and type of RTAs at a given point on the time line and the cumulative total at that position.

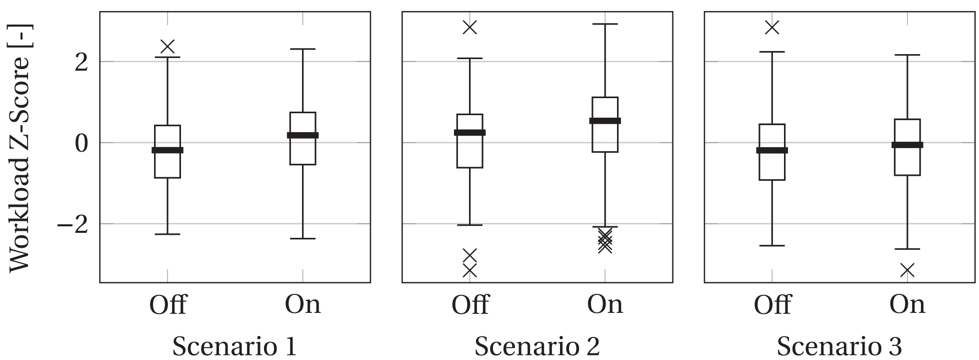


Figure 5.15: Normalised workload.

5.6 Discussion

The deviation from optimal total performance over all scenarios decreases when showing the deviation indicator. The difference is a non-significant effect, however. The central hypothesis—that using the deviation indicator would reduce the total performance deviation—can not be proven. The results from the experiment further demonstrate a substantial variation between subjects and also between scenarios.

Many participants appear to have used the two individual components of the deviation indicator, ATD and URT, to control one of these components rather than focus on overall performance. This focus on one performance element may have caused a lack of focus on the other element.

As with the experiments in Chapter 4, a considerable variation in performance is seen. Several metrics developed for this experiment—such as the amount ATD and URT—could be used to provide more directed training in working towards the planning objective.

The participants indicated that the deviation indicator mainly shows where deviations from performance occur. Rather than focusing on the total performance shown on the right side of the display, participants looked at the locations along the time axis where the deviations increased. The orange and green lines in Figure 5.1 indicate locations of potential optimisation. The total deviation, i.e., the sum of ATD and URT, is an unknown parameter to novice users of the new display. The acceptable vertical location of the total indicator has to be learned through experience. Secondly, the optimal total deviation depends on the actual traffic situation, as bunches of traffic will lead to—and ultimately require—much more deviation than a regularly spaced flow of flights.

Especially when considering ATD performance, the spread between participants decreases when the deviation indicator is available. In particular, the results show a reduction of the extremes. It appears that those participants mainly benefited from the visualisation of potential performance improvements. As such, the display could support inexperienced users in learning to recognise those opportunities for optimisation.

The three scenarios were constructed to provide comparable complexity and required performance. Each scenario required three distinct periods in which the optimal solution required using the secondary runway for an equal number of flights. In the experiments in Chapter 4, large differences in performance remained. In this experiment, the differences are not as pronounced in terms of spacing performance. However, considerable differences still exist in the two new performance criteria, ATD and URT. These differences in scenario complexity may well still have generated a confound in the experiment results.

A possible explanation for variation in complexity is found in the timing of the traffic: the three waves of high traffic load—for which the secondary runway was required—were timed differently in each scenario to avoid pattern recognition. However, the difference in timing may have contributed to the differences in performance since participants could have reacted to the difference in time between waves (i.e., intervals might have been too short or too long). As with the experiments in Chapter 4, observations during the experiment revealed that participants did not recognise scenarios at all, not even the training scenarios, which were repeated four times in total. This lack of recognition suggests that the variable timing of high traffic load was not necessary.

The current display concept is based on single use of the runways only. Departing aircraft are currently modelled as the time interval that the runway is not reserved. This assumption has two limitations. First, a runway can only be used for a departure when sufficient time is available for the aircraft to enter the runway and take off. Changing the configuration, therefore, is only meaningful when the change provides a natural multiple of such departure intervals. Secondly, many airports use runways for departures and arrivals simultaneously. Both limitations may be addressed by explicitly visualising departures on the time line. A number of the current operational systems already provide such arrival/departure support, which could be included in calculating the balance between demand and capacity.

The current demand and capacity model assumes that one second of delay is equally unfavourable to one second of arriving earlier and to one second of departure delay. However, the presentation using a line could work with more detailed models, as long as the contributions of ATD and URT can be meaningfully cumulated. Future work could investigate developing a more appropriate model.

Finally, the proposed display does not support the ATCO in choosing between different runway combinations. Especially at airports such as Amsterdam airport, many different combinations are available. Moreover, the selection is based not only on demand but also weather, runway availability, and legislative restrictions. To provide full support, such information should be included in the display.

5.6.1 Recommendations

Future research should focus on whether the observed decrease in variations due to the presence of decision support presented in this chapter can be confirmed. A similar human-in-the-loop experiment should be performed with more equal scenarios that put a lower workload on the controller. To that purpose, scenarios could be constructed with a better spaced initial arrival stream and run at a speed closer to real-time.

A larger number of participants with preferably a higher level of experience would provide more accurate and useful results, especially when a clear strategy for using this type of interface is explained beforehand. The choice of whether to explain the strategy and measure performance differences or not to explain a strategy and look for strategy differences should be made explicitly to obtain the desired outcome. When repeating the experiment described in this chapter, a clear strategy could be provided to the participants beforehand in order to increase their understanding of the task.

Finally, attention should be given to support of the runway assignment process, which is expected to result in an improved runway planning quality. For example, a future interface could provide a suggested assignment of aircraft to the available runways, which the controller only needs to accept if they agree. In this way, the controller's focus can be directed to optimising the runway planning.

5.7 Bibliography

[1] "Performance review report 2012," Performance Review Commission, EUROCONTROL, Brussels, Belgium, Tech. Rep. May, 2012.

[2] N. Hasevoets and P. Conroy, "AMAN status review 2010," EUROCONTROL, Brussels, Belgium, Tech. Rep., 2010.

[3] ——, "Arrival Manager - implementation guidelines and lessons learned," EUROCONTROL, Brussels, Belgium, Tech. Rep., 2010.

[4] Boeing, *Airport noise and emissions regulations*, 2012. [Online]. Available: `http://www.boeing.com/commercial/noise/` (visited on 04/16/2014).

[5] *Bewoners aanspreekpunt Schiphol*, 2014. [Online]. Available: `http://www.bezoekbas.nl/`.

[6] V. Kapp and M. Hripane, "Improving TMA sequencing process: Innovative integration of AMAN constraints in controllers environment," in *IEEE/AIAA 27th Digital Avionics Systems Conference*, St. Paul, MN, USA: IEEE, Oct. 2008, pp. 3.D.1–1–3.D.1–9.

[7] L. Boursier, B. Favennec, E. Hoffman, L. Rognin, F. Vergne, and K. Zeghal, "Combining sequencing tool and spacing instructions to enhance the management of arrival flows of aircraft," in *AIAA 5th Aviation, Technology, Integration, and Operations Conference (ATIO)*, Reston, VA, USA: AIAA, Sep. 2005, pp. 1–14.

[8] M. Schultz, H. Fricke, T. Kunze, *et al.*, "Uncertainty handling and trajectory synchronization for the automated arrival management," in *Second SESAR Innovation Days*, Braunschweig, Germany, 2012.

[9] H. Helmke, R. Hann, M. Uebbing-Rumke, D. Mueller, and D. Wittkowski, "Time-based arrival management for dual threshold operation and continuous descent approaches," in *8th USA/Europe Air Traffic Management Research and Development Seminar*, vol. 29, Napa, CA, USA, 2009.

[10] K. van de Merwe, E. Oprins, F. Eriksson, and A. van der Plaat, “The influence of automation support on performance, workload, and situation awareness of air traffic controllers,” *The International Journal of Aviation Psychology*, vol. 22, no. 2, pp. 120–143, Apr. 2012.

[11] AviBit, *Optamos arrival management system*, Graz, Austria, 2012.

[12] W. Nedell, H. Erzberger, and F. Neuman, “The Traffic Management Advisor,” in *Proceedings of the American Control Conference*, San Jose, CA, USA: IEEE, May 1990, pp. 514–520.

[13] H. Erzberger, “Design principles and algorithms for automated air traffic control,” in *AGARD Lecture Series No. 200 on Knowledge-based Functions in Aerospace Systems*, 1995, pp. 16–17.

[14] D. G. Denery and H. Erzberger, “The Center-TRACON Automation System: Simulation and field testing,” NASA Ames Research Center, Moffet Field, California, Tech. Rep. August, 1995.

[15] H. N. Swenson, T. Hoang, S. Engelland, *et al.*, “Design and operational evaluation of the traffic management advisor at the Fort Worth Air Route Traffic Control Center,” in *1st USA/Europe Air Traffic Management Research and Development Seminar*, Saclay, France, 1997.

[16] M. Tielrooij, C. Borst, D. Nieuwenhuisen, and M. Mulder, “Supporting arrival management decisions by visualising uncertainty,” in *Proceedings of the SESAR Innovation Days*, D. Schaefer, Ed., Stockholm, Sweden, 2013.

[17] A. J. Tattersall and P. S. Foord, “An experimental evaluation of instantaneous self-assessment as a measure of workload.,” *Ergonomics*, vol. 39, no. 5, pp. 740–8, May 1996.

[18] F. Wilcoxon, “Individual comparisons by ranking methods,” *Biometrics Bulletin*, vol. 1, no. 6, pp. 80–83, 1945. [Online]. Available: `http://www.jstor.org/stable/3001968`.

Chapter 6

Synthesis

The proposed displays in the last two chapters were demonstrated using a theoretical uncertainty model. This chapter applies the real-life uncertainty from Chapter 3 to evaluate the application of the display in actual operation using data from actual flights. Using the visualisation, this chapter then explores improvements in the uncertainty model and the diagram to support working with uncertainty in arrival management.

6.1 Introduction

Chapter 2 reviewed the current arrival management operation, its systems and its Human-Machine Interfaces (HMIs). A key finding was the presence and effects of prediction uncertainty on the process. The chapter proposed a future approach in which uncertainty would form an input to the decision-making process. Chapter 3 developed a method to predict the uncertainty of the arrival time using the available knowledge of a flight's progress. To use this information, Chapter 4 proposed an HMI that visualises the uncertainty in arrival time and its effect on the expected demand on a runway. Chapter 5 then extended the display to support decisions on opening a secondary runway and introduced a metric and visual presentation of the desired balance between demand and capacity.

Both Chapter 4 and 5 use a simulated arrival time uncertainty. The uncertainty was modelled as a normal distribution with a standard deviation that decreased linearly with decreasing prediction horizon. The chapters already recognise the difference between this distribution and the actual distributions found in flight data as described in Chapter 3.

This chapter aims to combine the predicted real-life uncertainty with the proposed display. By applying this combination, this chapter will evaluate the findings from the four previous chapters; It will discuss the suitability of the modelled uncertainty for operation and the suitability of the proposed display using that modelled uncertainty.

This chapter will use a theoretical schedule of flights with their flight state matching the real-life distributions in the dataset. The next section describes the theoretical schedule. Section 6.3 evaluates the presentation of real-life uncertainty from Chapter 3 on the proposed display. Section 6.4 then evaluates the validity and meaning of the resulting occupancy expectation under actual uncertainty. Finally, Section 6.5 discusses these results.

6.2 A hypothetical schedule

To evaluate the display using the predicted uncertainty from Chapter 3 we could use a sample of the collected Flight Update Message (FUM) data used in that study. As this is a dataset from actual operations, it contains many types of disturbances due to peaks in flight schedules, traffic mix variations, and missing data. Furthermore, Amsterdam rarely has a period of two hours in which a single runway is used except during the low traffic volumes at night. Therefore, this chapter uses a fictitious but representative sample with its principal properties derived from the actual data.

This schedule is used throughout the chapter to visualise the effects of potential modifications. The schedule assumes that an aircraft is predicted to arrive every 2.5 minutes with an occupancy interval of two minutes for each aircraft. This interval provides a 30-second buffer if all aircraft arrive at their predicted Estimated Time of Arrival (ETA). While synthetic and by no means reaching the runway capacity at Amsterdam, this allows for a more straightforward evaluation of the display given the high uncertainties found in the FUM data.

The flight state is randomly taken from the distribution of flight states from the data in Chapter 3 at a given point on the prediction horizon, as indicated in Figure 6.1. This figure shows the fractional distribution of flight states at a given prediction horizon (i.e., the time line presentation's horizontal axis). As with all figures in this chapter, the colours match their flight state.

The distribution of states of flights that are predicted to land immediately (i.e., 0 minutes on the horizontal axis) best explains this figure:

- 84% of flights have a confirmed airborne (AA) state.
- 13% of flights are assumed to be airborne (TA). Those flights are not yet about to arrive, but their arrival time estimate has not been updated.
- Finally, 3% of flights have other states (FI, SI, or Other). No information has been received for these flights, and, most likely, these flights have not departed.

As the prediction horizon increases (i.e., further to the right along the horizontal axis), a lower fraction is confirmed to be airborne and replaced by other states. Most notably, only a flight plan is available (FI) for an increasing share of the flights.

Figure 6.2 shows what the proposed interface would look like according to the theoretical schedule. This time line uses the Johnson distributions from Chapter 3. The distribution of the flight states against the horizon (the horizontal axis) is as expected: many airborne (AA) flights close to arrival (green) and many flights for which we only have a flight plan at the right-hand end of the time line (FI, purple). The first eight flights have a predicted arrival within 20 minutes. Since the analysis in Chapter 3 only starts at a prediction horizon of 20 minutes, these flights do not have an estimated uncertainty. The first two flights that are expected to arrive have flight state TA (light blue). In reality, it is unlikely that those flights will arrive at the indicated time, but as their planned departure time has long passed, EUROCONTROL has assumed these to be airborne.

The occupancy indicator of the later flights shows no gaps between flights. The ETAs of the flights are spaced regularly and at intervals larger than the hypothetical two-minute spacing requirement. According to the prediction, the flights would fit on the runway, but the uncertainty in arrival time leads to a possibility that minimum spacing could be violated. The variation even leads to a locally predicted excess of demand around 12:50.

The ETAs in Figure 6.2 represent the reported ETAs. The Probability Density Functions (PDFs) of the error distributions have a bias, mostly toward a delay.

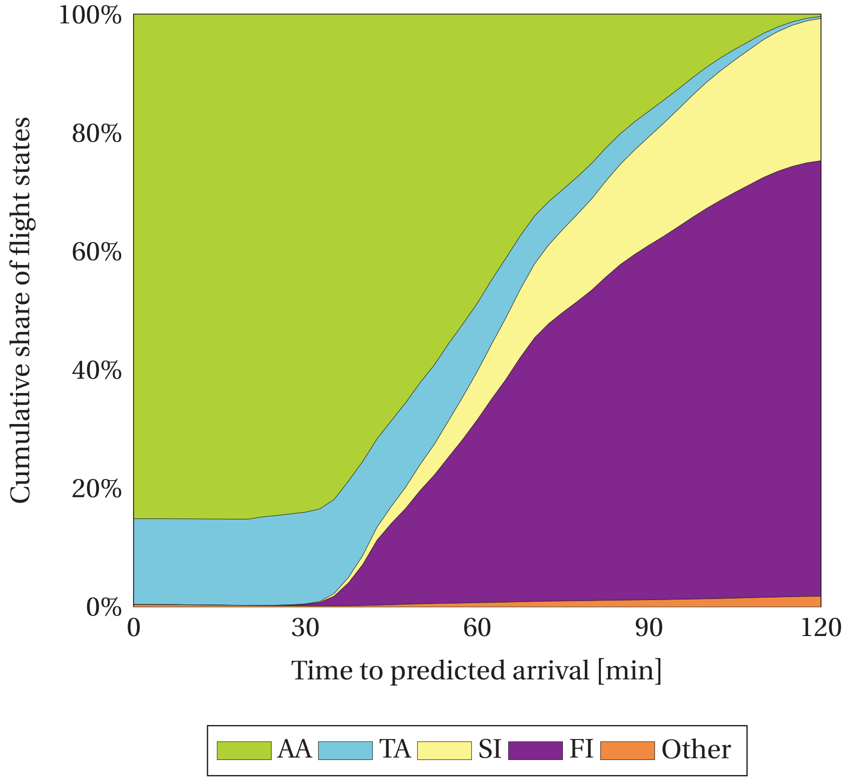


Figure 6.1: Distribution of flight states as a function of predicted time to fly (the prediction horizon).

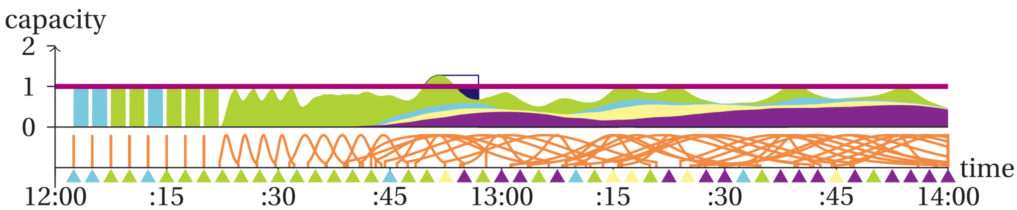


Figure 6.2: Hypothetical time line with uncertain ETA based on the FUM data.

It would be possible to adjust the location of the triangles to the median of the distributions to show the most likely arrival time. This would in effect “correct” the ETA for the most likely error. As this work is focused on the uncertainty of the ETA, such a correction is out of the scope of this work.

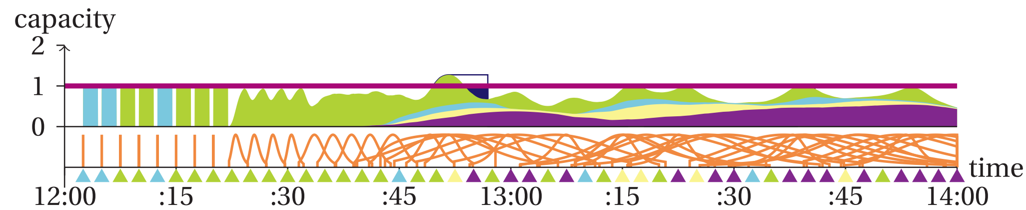
6.3 Presentation of uncertainty

The analysis of uncertainty in Chapter 3 shows that, despite modern trajectory prediction capabilities, the current aviation system is not yet able to provide ETAs to an accuracy at which it is possible to define a sequence for an entire two-hour horizon. The ambiguity in the sequence is visible in the time line of the hypothetical scenario (see Figure 6.2). This section will further investigate this finding and discuss methods to reduce that ambiguity.

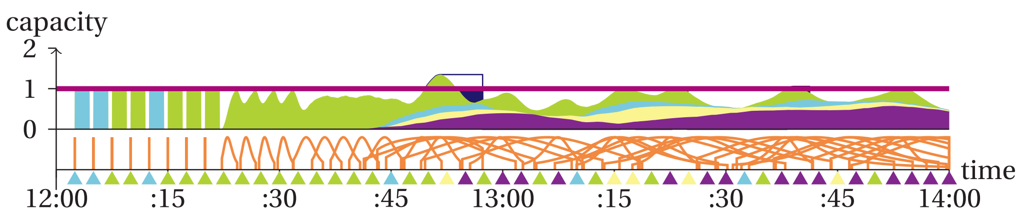
Section 6.3.1 discusses the amount of uncertainty shown on the time line. The next section addresses the variation in the uncertainty. In Section 6.3.3 an initial exploration is done to investigate the possible effects of the use of Airport Collaborative Decision Making (A-CDM) data on the uncertainty.

6.3.1 Amount of uncertainty

The PDFs of multiple flights overlap, especially at a larger prediction horizon. This overlap implies that those flights can theoretically have a trajectory that would make them arrive simultaneously if Air Traffic Control (ATC) would not separate them. In the fictional time line, all PDFs of the flights due to arrive within 20 minutes or more overlap with at least one other PDF. At the rightmost end of the time line, two hours before arrival, a PDF for a flight in FI state overlaps about 20 others. Showing individual PDFs becomes problematic, as the overlap causes clutter on the proposed diagram. The uncertainty level is too large for an operator—or an automated system—to be helpful.



(a) Uncertainty limited to the central 95% of the distribution.



(b) Uncertainty limited to the central 90% of the distribution.

Figure 6.3: Effects of applying a smaller confidence interval.

One of the options to address the overlap is limiting the minimum likelihood at which the PDF is shown. One could argue that the flight's potential arrival below a certain threshold becomes operationally irrelevant. Whereas, Figure 6.3(a) has been limited to 95% of the spread of the PDF, Figure 6.3(b) uses a likelihood of

90%. As expected, more narrow peaks are visible, and the peaks also become more pronounced. The predicted excesses are more clearly defined.

6.3.2 Variation in uncertainty

For many flights in the two-hour horizon, the only available information is the filed flight plan (FI). The effect of this lack of information can be demonstrated by comparing two flights in Figure 6.4. The figure shows that the ETA of a flight with a Filed (FI) state (the left purple triangle) is at 13:00. However, its PDF indicates a 95%-likelihood of arrival between 12:44 and 13:12. The (purple) occupancy expectation starts at the start of the PDF and shows that the flight could occupy the runway anywhere from 12:44 until 13:14. This interval gives an uncertainty of 30 minutes at 60 minutes before arrival.

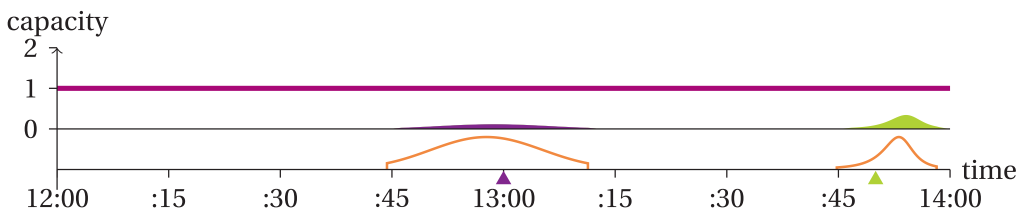


Figure 6.4: Two flights on the fictional time line with different Johnson distributions.

Simultaneously, the uncertainty of some airborne flights is much smaller, even at a long horizon. A flight with an airborne state (AA) at 13:50 (the green triangle on the right), 100 minutes from the current time, has an 95% likelihood of arriving in an interval of about 13 minutes. The occupancy expectation runs from 13:45 until 14:00. If no other arrivals would be expected around the expected landing time of such an accurately known flight, its landing could be scheduled, even if earlier arrivals could not be. This timely information would allow a predictable operation for that particular flight.

These two examples demonstrate the considerable variation in uncertainties that can exist on the display at a single moment in time. In this example, any action on the first aircraft would be sensitive to corrections. However, this fact alone does not prevent the display from being practical. For example, a reasonable operational scenario would be the end of an evening where the last flights of a busy period will need planning on a short horizon. However, ATC can already plan later flights in the more quiet night.

6.3.3 Reducing uncertainty and variation through A-CDM

Further reduction of the uncertainty would enable planning at a further horizon. As the previous sections demonstrate, many flights have no further information than a flight plan. Figure 3.12 in Section 3.4 shows that the shape of the distribution in the Filed (FI) state does not change significantly with the prediction horizon. Chapter 3 explains that the factors that introduce uncertainty follow from events rather than gradual disturbances. As long as we cannot predict those events, uncertainty will not become smaller. The FUM data do not capture events that lead to deviations in the Filed state.

Chapter 3 suggests the use of Airport Collaborative Decision Making (A-CDM) information as a source. A-CDM information in the dataset shows potential benefits of using such additional information in two ways: Earlier information on the—unconfirmed—departure of the flight and more accurate predictions of the arrival times of flights with such an unconfirmed departure state (TA). This subsection will perform an initial exploration of the effects of this data using the time line and our fictitious schedule. Note that this section only uses the knowledge that A-CDM information is available to the EUROCONTROL Network Manager (NM).

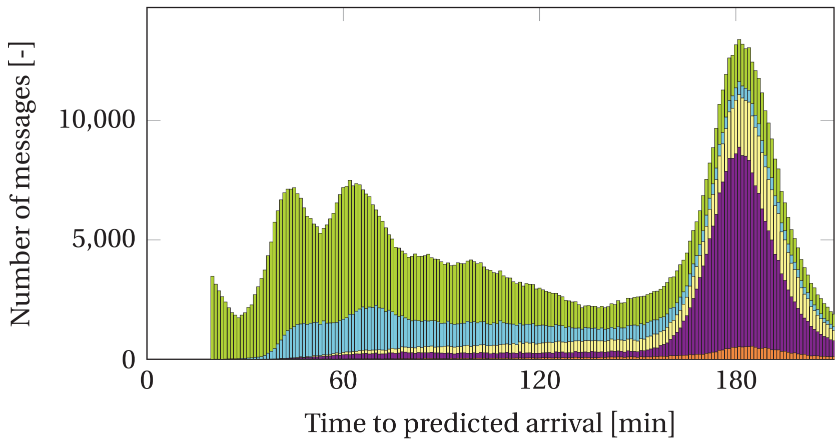
The first effect is visible in Figure 6.5. This version of the analysis—shown earlier in Figure 3.5—separates flights based on the availability of A-CDM data for those flights. The graph shows a much larger fraction of flights that are assumed to be airborne (TA). These messages provide a supplementary estimate of the ETA between the filed (FI) state and the airborne (AA) state.

The different distribution of messages is also visible in the distribution of flight states against horizon, as seen in Figure 6.6. A much smaller fraction of flights from airports that provide A-CDM information has a FI state at a given horizon. Figure 6.7 shows the same distribution of flight states as Figure 6.1, but a much lower fraction of flights has an FI state at each horizon. The most likely explanations for this difference are that A-CDM data allows the NM to predict a flight's time of departure better, and that the message set contains an explicit message confirming take-off [1]. NM can, therefore, more confidently assume that a flight is airborne without radar confirmation.

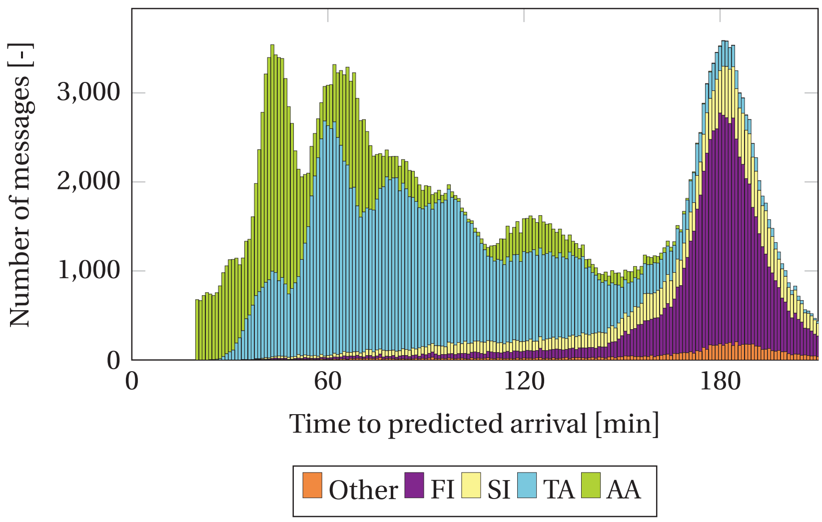
Note, however, that at horizons beyond 45 minutes, the share of flights with an AA state is lower for flights from an A-CDM airport. This is because the increased share of confirmed airborne flights includes the flights that depart beyond the two-hour horizon and have had more time to be detected as airborne.

A second effect of the A-CDM information is that the TA state estimates are more precise than those for the FI state. Also, these estimates are less uncertain than those for flights with a TA state but do not come with A-CDM information. This difference aligns with the objective of A-CDM to improve departure predictability and, hence, predictability of the whole flight.

The difference in accuracy can be demonstrated by comparing the distributions for both sets of flights. Figure 6.8 provides an initial analysis of the distributions. The graph shows that the uncertainty for flights with A-CDM information is smaller, particularly at the left-hand side of the distribution representing flights that arrive



(a) Message count versus predicted time to fly for flights without A-CDM data.



(b) Message count versus predicted time to fly for flights with A-CDM data.

Figure 6.5: Distribution of message types with and without A-CDM information. Note that the number of flights with A-CDM information in the bottom plot is much smaller—hence the lower total.

earlier than predicted.

The fact that flights originate from an A-CDM airport affect the uncertainty in two ways: less flights with FI state in the left half of the time line and less uncertain predictions for the TA state. Figure 6.9 shows the effect of the double-edged improvement of accuracy by generating the same time line using the distribution of flight states according to Figure 6.7. The first notable difference is the larger number of flights with a TA state (blue) and fewer flights with an FI state (green). Secondly, the lack of flights with an FI state around the 60-minute mark decreases

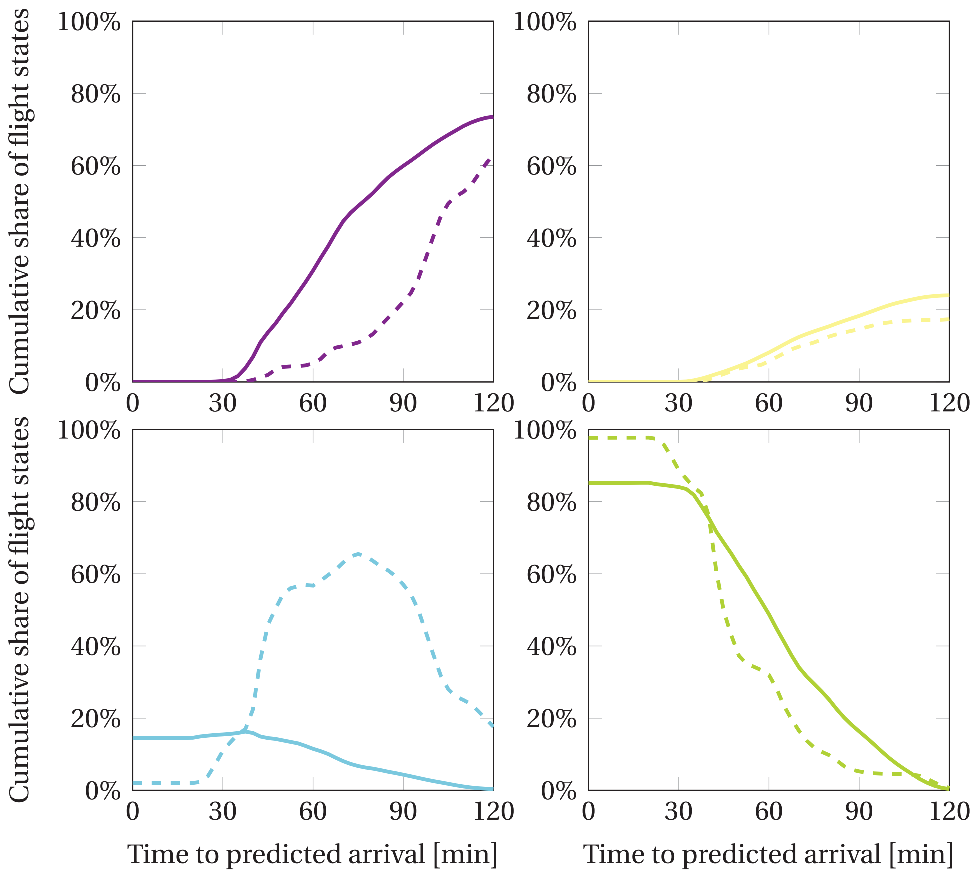


Figure 6.6: Effect of A-CDM on share of flight states versus prediction horizon. The dashed line is the share of flights with A-CDM information. The colours correspond to the flight states in the previous graph.

uncertainty, leading to more distinct peaks and reduced expected excess demand.

The data in this analysis are from 2013 and 2014. Two years after this study—in 2016—the number of connected airports in Europe was 18 [2, p. 3]. That number has risen to over 54 in 2020 [3, p. 31], allowing for a far more detailed analysis. Furthermore, this initial exploration did not explore the information in the A-CDM status. Repeating this analysis based on current data and using the A-CDM states themselves would be worthwhile.

Even more systems provide estimates or information on the flight's arrival time during a flight's progress. The last components in that chain are the Trajectory Predictors (TPs) in the AMAN itself. Besides providing more extensive information in a single data source, these other sources could be combined. Such an approach could provide the ETA with the associated uncertainty of the most accurate source as these become available.

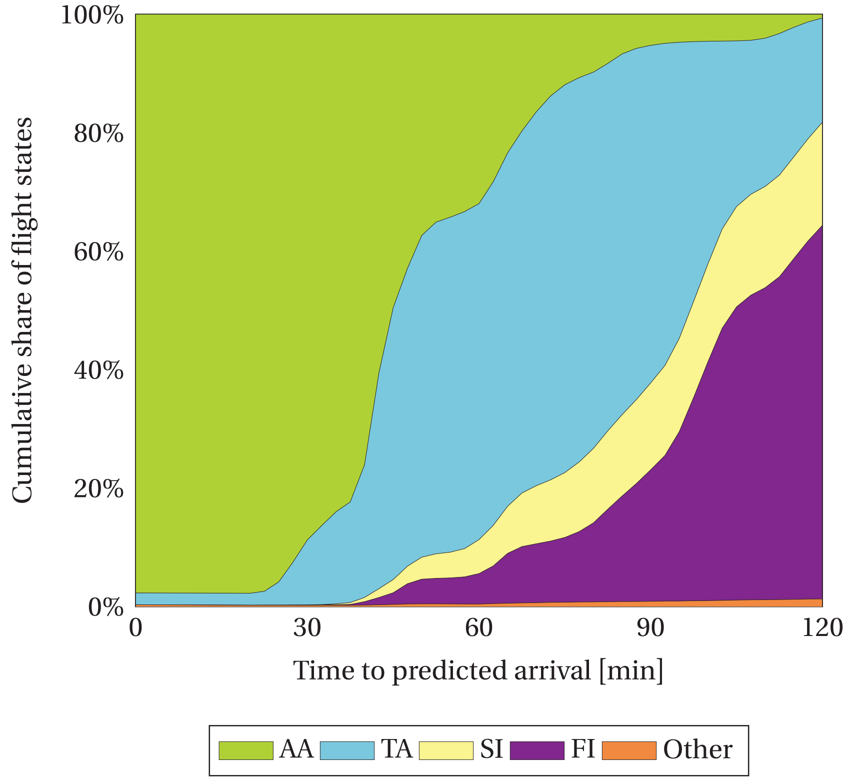


Figure 6.7: Distribution of flight states as function of prediction horizon with A-CDM data.

6.3.4 The limits of uncertainty

If an arrival management process is to support a high-capacity operation, the intervals between consecutive flights should be as close to the required minimum as possible. In such a full schedule, a planning error would require re-planning many flights to accommodate an incorrectly planned flight.

The above analysis shows that much uncertainty continues to exist, and a considerable component of that uncertainty is based on events *before* departure. Prediction capabilities can certainly still improve, as they have done in the past. Data sharing, such as A-CDM, may support more information for the prediction as events unfold. However, preventing those events would require removing uncertainties from the operation.

As recognised in Chapter 3, the second route involves improving our ability to work with uncertainty when possible. Even if planning on the full extent of the time line is not possible, the display concept may support dynamic adjustment of the horizon.

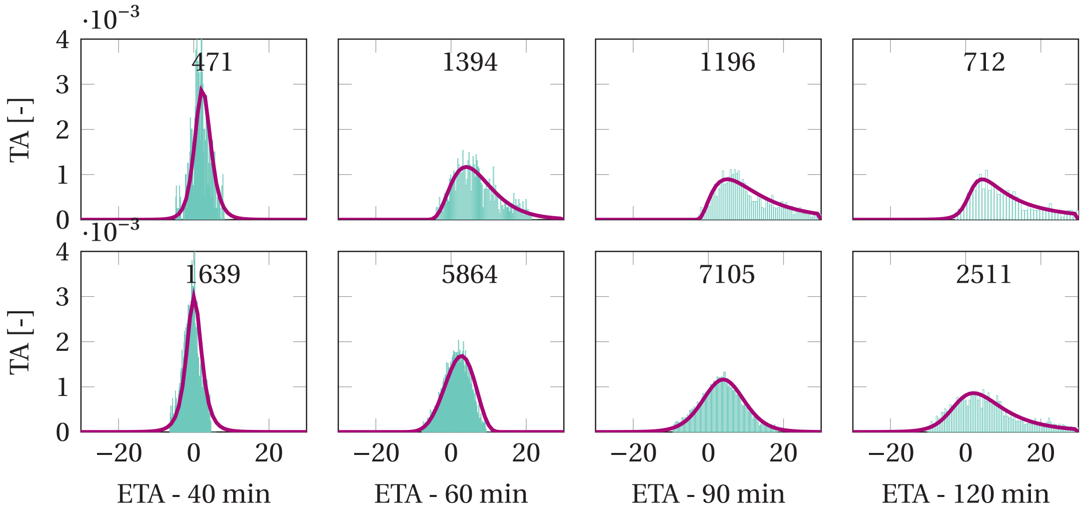
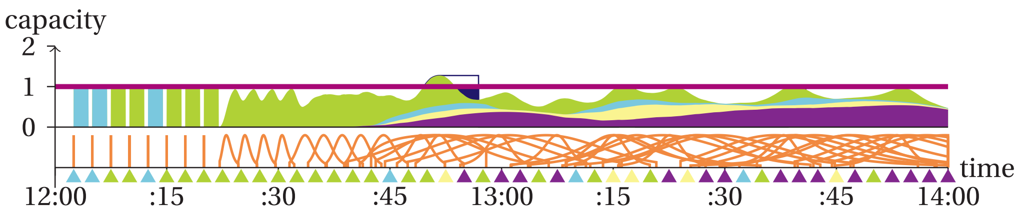
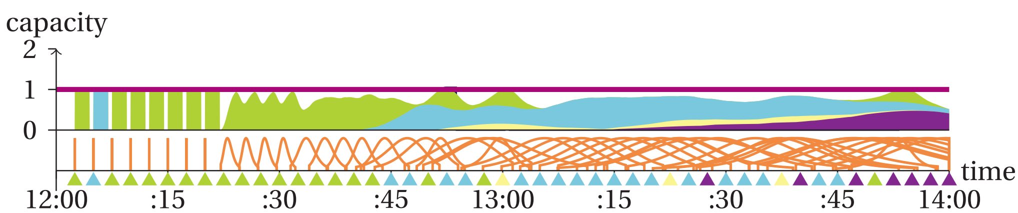


Figure 6.8: Difference in uncertainty between flights with (top) and without (bottom) A-CDM information.



(a) Fictional time line for flights, without A-CDM information.



(b) Fictional time line for the same flights, with A-CDM information.

Figure 6.9: Effect of A-CDM data on the fictional schedule.

6.3.5 Working with uncertainty

Uncertainty increases with the prediction horizon. As shown in the hypothetical time line in Figure 6.1, at some point on the horizon, the PDFs overlap and become too wide to be of use to an operator. Beyond that point on the horizon, the hypothetical use of the display for planning a sequence would not be practicable. The exact point at which this is possible depends on the local uncertainty of the sequence. The uncertainty will change depending on the state of the flight and the time of day. This dependency on the horizon implies that a hypothetical suitable working horizon may change over time.

A human operator or automated system could benefit when uncertainty is low and apply caution when uncertainty is high when aware of the actual workable horizon. To do so, the arrival time distributions can provide the likelihood of a change in the sequence. A way of doing so is by determining the likelihood that a flight in a sequence will be overtaken.

For every flight, the probability of such a change in sequence is the complement of the probability that any of the flights behind it will not overtake that flight (see Equation (6.1)).

$$P_{first_i(t)} = P_i\left(t\right) \prod_{j=i+1}^{j=n} \left(1 - C_{j(t)}\right) \tag{6.1}$$

Here $P_{first_i(t)}$ is the probability for aircraft i at time t to arrive before any of its planned followers. This is a function of its own probability to arrive at that time $P_i\left(t\right)$ and the probability $C_{j(t)}$ that a follower j has arrived by that time.

The probability of a swap in sequence P_{swap_i}, is the complement of integrating Equation (6.1) over time. This integration represents the probability that aircraft i arrives before its planned followers regardless of the time it arrives (see Equation (6.2)).

$$P_{swap_i} = 1 - \int_{\infty}^{t=0} P_{first_i(t)} dt \tag{6.2}$$

Figure 6.10 shows the probability of a swap in sequence for every flight on the time line on the bottom graph. In this sequence, a swap is highly likely (0.8) from 12:45 onward. The cause of this high probability is the aircraft with a TA state (blue) that is likely to be delayed according to its distribution. This flight is followed by two flights with an AA state (green), which have lower uncertainty and less predicted delay.

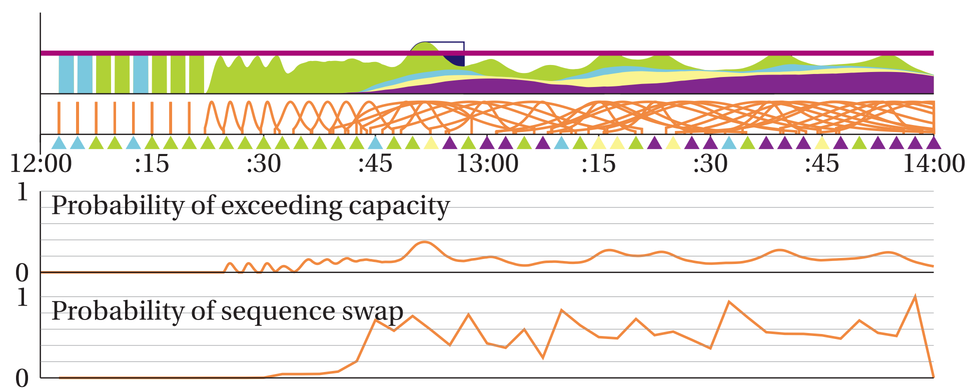


Figure 6.10: Probabilities of exceeding capacity and swap of position.

Note that Equation (6.1) and Figure 6.10 show the probability that each individual flight will be overtaken by its planned followers. It underestimates the likelihood of any sequence swap before that point in the sequence. The total *running* probability can then be approximated by the product of the inverse of that graph. This product represents the cumulative probability that a swap will have occurred at that point. The mentioned flight at 12:45 has a probability of a sequence swap of 0.8. The graph in Figure 6.10 then remains between 0.2 and 0.8. Therefore, the product will likely be very close to 1.0 soon after 12:45.

The graph in Figure 6.10, possibly combined with a running product, provides an example of a dynamic calculation of the limit of the planning horizon. Further research on this subject may be of value in both automatic decision systems and interfaces for human operators.

6.4 Presentation of occupancy

The proposed display aims to incorporate a visualisation of the mathematical concept of uncertainty into the operator's decision-making process. By integrating the uncertainty, the display also provides an indicator of the balance in demand and capacity. However, translating the mathematical concept into a usable display is far from trivial.

The width of the uncertainty is often more extensive than the width of the landing interval. Such high amounts of uncertainty introduce a problem with the meaning and validity of the occupancy expectation and, therefore, the resolution indicator. This problem follows from two aspects: the fact that the sequence is undefined and the fact that an expectation value seldom indicates certainty. Therefore, this section discusses the presentation and validity of the occupancy expectation.

6.4.1 Uncertainty of sequence

When the flights' PDFs overlap considerably, the aircraft sequence is unknown. In addition, the required interval between consecutive aircraft follows from their wake turbulence categories (see Section 2.2.2). Thus, if the required spacing depends on which one of two aircraft lands first, the actual required interval is unknown. Of course, this problem exacerbates when multiple aircraft have overlapping potential arrival times. As Figure 6.2 shows, such a scenario is a realistic possibility with the current level of uncertainty.

Solutions for this problem may be sought in defining the required interval using less precise means. An example would be to calculate an average separation for the traffic mix. This approach would form a dynamic bridge between the current concept of declared capacity—often in flights per hour and specified for an hour—and the actual spacing based on the final approach planning sequence.

6.4.2 Meaning of occupancy expectation

If the arrival time is known, the required capacity can be represented on a time line as a block. This approach from Chapter 4 allows visualising demand and capacity. By doing so, the diagram can support the operator in managing this balance. However, when a PDF represents the arrival time, the presentation of occupancy becomes an expectation value. An occupancy expectation exceeding the capacity does not necessarily imply that the runway will be occupied. Similarly, when the expectation does not exceed capacity, there is no guarantee that no conflict will occur.

The transition from a pure occupancy to an expectation also reduces the validity of the resolution indicator. Since the expectation assumes that a flight could arrive anywhere within its PDF, the occupancy distributes from the start of that PDF to its end plus the required interval. This spreading of the occupancy means that the planning problem can occur anywhere on that distribution.

To demonstrate, we can use the example of two overlapping PDFs wider than the runway occupancy interval for these aircraft. In that case, the maximum occupancy expectation of each of the aircraft will not reach half of the capacity. Therefore, the combined occupancy expectation would never exceed the runway's capacity. Since demand does not exceed capacity, the resolution indicator would not appear. However, as the PDFs overlap, the aircraft could arrive simultaneously, which means that the potential conflict only resolves when both landing intervals have passed. This mechanism makes the resolution indicator an expectation as well.

Using the occupancy expectation of each flight, it is possible to calculate the probability that demand exceeds capacity: The occupancy expectation represents the probability that an aircraft will occupy the runway at a given time. For one runway, the probability of exceeding the capacity is equal to the probability of more than one aircraft occupying the runway simultaneously. Equation (6.3) calculates this probability by taking the complement of the probability that either none or exactly one aircraft occupies the runway:

$$P_{O>1(t)} = 1 - P_{0aircraft(t)} - P_{1aircraft(t)}. \tag{6.3}$$

The probability of zero aircraft is equal to the product of the probability that no flights have arrived yet:

$$P_{0aircraft(t)} = \prod_{n}^{i=0} \left(1 - O_i\left(t\right)\right) \tag{6.4}$$

The probability of exactly one aircraft is equal to the sum of all flights' probabilities to be the only one occupying the runway at that time. The latter is equal to the occupancy expectation of that aircraft times the complement of the occupancy expectations of all others:

$$P_{1aircraft(t)} = \sum_{n}^{i=1} O_i(t) \prod_{n}^{j=1,i\neq i} \left(1 - O_i(t)\right) \tag{6.5}$$

Figure 6.10 shows this probability as the flight graph under the time line. As the graph shows, demand may exceed capacity from 20 minutes onward. The probability of a conflict remains at about 20% over the entire remaining line, with a peak of about 40%. The latter peak matches the point when the occupancy indicator also shows an excess in demand. This peak shows that—while the occupancy expectation is greater than capacity—it is more likely than not that no conflict will occur.

Both the occupancy indicator and the resolution indicator become expectations in a display that uses uncertainty. More research is needed to determine whether operators can understand and use the indicators, knowing that these are expectations. The graphs in Figure 6.10 may well help with this.

6.5 Discussion and conclusion

This chapter combines the prediction of uncertainty from Chapter 3 with the display developed in Chapters 4 and 5. By doing so, it aims to validate the display in an actual operational context.

6.5.1 Suitability of the display concept

Implementing actual uncertainty derived from the data from 2013 strongly reduces the visual meaning of the presentation through the occupancy indicator. Two properties of the real-life flight data are responsible for this reduction: The amount of uncertainty and the variation in that amount.

The uncertainty in the ETA is much larger for most flights than a typical landing interval. Therefore the PDFs of these flights overlap. The individual contribution of a flight can no longer be distinguished from the diagram. The variation of uncertainty between different flights is quite considerable. Even if a particular flight has a small uncertainty, it is likely to overlap with a flight with very high uncertainty. This reduces the value of any planning action on a predictable flight.

The high number of aircraft that have an overlapping PDF also reduces the meaning of the occupancy indicator. A visualised excess of expected occupancy or a lack of excess does not necessarily indicate the probability thereof. The related delay indicator follows from the visualised occupancy and has the same problem.

6.5.2 Potential improvements

Several routes for improvement of the concept have been explored. These routes address both the amount of uncertainty and the visualisation and use of the information.

A significant contribution to the amount and variation of uncertainty is due to aircraft that are still on the ground. This matches the observations in Chapter 3. The flight state in the FUM data provides no information on the progress toward departure. Preliminary use of A-CDM information in the recorded dataset reduces the uncertainty at the intermediate horizon. Since far more airports are now providing A-CDM information, the benefit of this approach is likely to have improved.

While the occupancy expectation is not a complete representation of the likelihood of a planning conflict, the underlying probabilistic information does allow the exact calculation of that probability. This probability is no longer a visual representation of the balance between demand and capacity. However, it can help in deciding on the need for corrective action.

The uncertainty varies per flight but typically increases with the prediction horizon. The prediction horizon at which the use of the display becomes infeasible will, therefore, change over time. It is possible to calculate the probability of a planning error using information on the uncertainty. This calculation could allow a dynamic decision horizon based on the nature of the traffic situation. Such an approach could fit into a concept where, when nearby a departure is about to leave, the planning horizon is—temporally—adjusted to cope with that flight's high uncertainty [4].

6.6 Bibliography

[1] EUROCONTROL, "Airport Collaborative Decision-Making (A-CDM) implementation manual," EUROCONTROL, Brussels, Belgium, Tech. Rep., 2017.

[2] "A-CDM Impact Assessment," EUROCONTROL, Tech. Rep. March, 2016.

[3] "Network operations report 2020," EUROCONTROL, Tech. Rep., 2020.

[4] A. Vanwelsenaere, J. Ellerbroek, J. M. Hoekstra, and E. Westerveld, "Analysis on the impact of pop-up flight occurrence when extending the arrival management horizon," in *12th USA/Europe Air Traffic Management Research and Development Seminar*, Seattle, WA, USA, Jun. 2017.

Chapter 7

Conclusions

This chapter reflects on the material presented in the previous chapters by combining their results and conclusions. In doing so, it aims to develop an overarching view of the requirements for arrival management systems, how they deal with prediction uncertainty and how relevant information can be displayed. The chapter also discusses the difficulties of evaluating novel interfaces designed for an operation that does not exist yet.

7.1 Expanding the arrival management horizon

Current arrival management concepts for capacity-constrained airports are mostly limited to the Area of Responsibility (AOR) of each individual Air Navigation Service Provider (ANSP). Future operational concepts, such as those that are proposed in Single European Sky ATM Research (SESAR) and Next Generation Air Traffic Management System (NextGen), foresee an increase in the planning horizon to improve the long-term efficiency and predictability of operations. Depending on the concept, future horizons are expected to be at 200-500 NM from the airport, well beyond one hour of flying time [1], [2]. In the European Union (EU) the expansion to 180 NM from the airport has recently (2021) been codified in law through Common Project One [3].

Currently, the horizons at which Arrival Managers (AMANs) are used to monitor and influence traffic are to typically 20 to 30 minutes from the airport—or 150-200 NM flight distance [4]. These present-day horizons are limited because of three factors:

1. The availability of information on the predicted arrival time of aircraft, due to the geographical limits of surveillance,
2. the authority to influence the flight's trajectory, currently limited by the AOR of the relevant ANSP, and
3. the reliability of the predicted arrival times.

The first two limitations are to be addressed through sharing information on the flight's progress and constraints on arrival time set by the destination ANSP through System Wide Information Management (SWIM) [3], [5]–[7]. The uncertainty in a predicted arrival time limits the horizon at which decisions on runway assignment, routes, sequences and landing times remain effective. Earlier decisions have a higher chance of requiring revisions downstream, reducing the efficiency of the flight overall.

Finally, these concepts see a role for the human controller as *automation manager*, not the *automation monitor* [5, pp. 45]. If a controller is made responsible for arrival management, they will need suitable tooling to monitor and control that process. When the horizon increases, more processes and constraints will become part of that operator's work domain. The current Human-Machine Interfaces (HMIs) provide limited information on the constraints of the processes or the degree to which objectives are met.

This thesis aimed to develop such a (semi-automated) decision support for air traffic controllers performing the arrival management task under prediction uncertainty. To do so, it tried to achieve two objectives (Chapter 1):

- To develop a method to determine the uncertainty associated to an Estimated Time of Arrival (ETA).
- To develop an HMI that allows the operator to perform arrival management in the presence of uncertainty in the ETAs.

This chapter first discusses the identified impact of uncertainty on the arrival management process and the proposed method to determine that uncertainty during operations. Secondly, it describes the development of the HMI by evaluation of the current HMIs, the proposed visualisation of uncertainty on such an HMI, the results of testing the display in human-in-the-loop experiments and a discussion on the suitability of such a visualisation in real-life operation. Finally, the chapter concludes with recommendations and conclusions, looking back at the study.

7.2 Uncertainty as information

Key to arrival management is the ability to predict the arrival time of a flight accurately. Such trajectory prediction (Section 2.4) takes a flight's current state and intent and models how the aircraft will progress to its destination. Section 2.4.7 described how assumptions are made at all steps of the prediction process. While necessary to make any prediction, any inaccuracy in assumptions increases the probability of an error in the arrival time. A plan based on an incorrectly predicted arrival time may require replanning—and associated adjustment of trajectories—at a later point in time. If the solution was optimal before re-planning, such an adjustment causes deviation from that optimum.

This section evaluates the effects of uncertainty on the working horizon. It then describes the proposed method to predict uncertainty during live operation. Finally, it discusses further potential for development of this technique.

7.2.1 Uncertainty and the effective working horizon

With an increased prediction horizon, the chance of prediction errors increases in two ways: errors in flight modelling have a longer time to affect the prediction, and more modelled processes become part of the flight's trajectory. The accuracy of the prediction of aircraft trajectories has improved considerably and is still improving (Section 2.4.7). In Europe, however, many flights are shorter than the prediction horizon (Section 2.2). While planning their arrival time, many aircraft will not have left their departure airport. At this point, trajectory prediction has to include manoeuvring on the ground and processes during boarding. The scope of processes to be modelled then rapidly increases, making it unlikely that complete modelling is possible.

To accommodate for prediction uncertainty, present-day arrival management operations mostly use a static horizon before which no planning is performed. This horizon is often determined through experience by evaluating at which point re-planning occurs too often. Since the horizon is static, it does not account for variation in uncertainty over time. However, such variation is likely, for example due to traffic patterns in which the fraction of long-haul (lower uncertainty) and short-haul flights (higher uncertainty) varies.

A static horizon may at times be too conservative when predictions are far more accurate. In such a case, the potential benefit of arrival management is not fully realised. At other times, for example during adverse weather conditions, the horizon may be far too optimistic. Then the necessary re-planning would cancel out potential benefits.

7.2.2 Predicting uncertainty

Since the uncertainty governs the working horizon, it is a crucial attribute in the work domain of arrival management. If such information is available, it may be used in risk-based approaches: Deciding on when to take action based on the risk of that action being counterproductive. This approach requires the uncertainty of the prediction to be known or estimated.

Several concepts have been developed that predict the uncertainty as part of the process of predicting the trajectory (Section 3.2). Such simultaneous calculation of uncertainty supports a tailored forecast of each prediction's potential error. However, these techniques come at higher computational costs per the trajectory prediction process. Furthermore, each user of uncertainty information will either need to receive the uncertainty information with the trajectory prediction or make its own trajectory prediction to get the uncertainty information.

Many studies have been performed on the accuracy of different trajectory predictors [8]–[11]. Chapter 3 turned this process around: By analysis of the accuracy of the prediction, we may be able to obtain an estimate of the uncertainty of that prediction. By doing so, *trajectory prediction* and *uncertainty prediction* are disconnected. Through this disconnection, the method also provides room for expansion based on the available information.

By analysing the errors in the Flight Update Messages (FUMs) it is possible to develop empirical error distributions for the prediction of arrival time, given the properties of the flights (Chapter 3). This estimate of flight progress from the EUROCONTROL Network Manager (NM) provides the flight plan and updates on the flight to the relevant ANSP, and it often yields the first available estimate on a flight's ETA. As such, it is a good example of an *external* prediction that is locally used as a source of information.

The data from 2013 and 2014 show that the shape of the error distribution mainly depends on:

- Flight state, describing whether the flight has taken off (or is assumed to have done so). This corresponds to the recognised problem of predicting trajectories for aircraft that are still on the ground.

- Prediction horizon. This corresponds to the duration that modelling errors can have an effect.

- Time of day. This factor most likely relates to the impact of traffic density and type of operation on potential deviations from the initial trajectory.

The versatile Johnson distribution [12] allows for capturing the differently shaped distributions using four parameters. By doing so, the variety in distributions may be tabulated to allow for fast online uncertainty estimation. Such an online technique will support use in real-time applications such as HMIs for air traffic controllers.

Calculating the probability of runway occupancy (Section **??**) allows for the calculation of two further metrics: Sequence stability, expressing the likelihood of a sequence swap (Section 6.3) and the probability of a planning conflict (Section 6.4.2). These metrics could support active adjustment of the arrival management horizon based on the uncertainty at that moment.

7.2.3 An expandable method

The demonstration of the empirical method is based on available data from 2013 and 2014. However, the approach is independent of the data source. This independence allows application of other data sources. Just the knowledge that a flight originates from an Airport Collaborative Decision Making (A-CDM)-enabled airport improves the estimated accuracy of predictions for that flight (Section 6.3). This is without exploring the differentiation in A-CDM-states provided in the FUM data.

The FUMs were selected for their availability. Currently (2022), the NM's ETFMS Flight Data (EFD) contains much richer information to which the same method could be applied. Through the implementation of SWIM, far more predictions will become available as more parties share their prediction of the arrival time. Finally, the method could also apply to an ANSP's own trajectory prediction system providing an early route toward using uncertainty information.

7.3 Displaying arrival management and uncertainty

In Chapters 4 and 5 a method was proposed to visualise the uncertainty in the presented arrival time and its impact on the plan and the operator's decision-making. By doing so, the ubiquitous AMAN time line was extended to provide more information on the arrival management work domain.

This section first discusses the current HMI and the identified shortcomings with regard to the operation. Addressing a number of those shortcomings allows for relating the uncertainty to the effects on the work domain. The section then describes the result of the human-in-the-loop experiments. The last part of this section reflects on the results of the experiments to discuss whether the task of an air traffic controller at long time horizons is not different from the task at the shorter horizon.

7.3.1 Supporting the human operator

The current HMIs all show a time line with the aircraft label connected to its planned arrival time. This display provides a good indication of the current plan but little of the constraints to which that plan is a solution. When, as is common at the moment, an automated system generates such an initial plan, the human operator has limited ability to follow the automation's rationale in this plan.

Using Rasmussen's Abstraction Hierarchy [13] Section 2.7 explored the work domain of arrival management and its representation on the current HMIs. The most notable missing attribute of the work domain is the required spacing between aircraft. While often resolved internally by an automated spacing algorithm, the constraint itself is not presented.

7.3.2 The work domain as a diagram

In Chapters 4 and 5, a diagram was developed using a method inspired by Ecological Interface Design (EID) to reveal the content and structure of the work domain. A fundamental basis of the diagram is the presentation of the required spacing as an area representing demand. By doing so, the higher level goals of demand and capacity can be related to the lower level spacing constraints between flights. In effect, the method merged the concept between two types of displays often found in flow control and arrival management: The AMAN time line and visualisations of demand and capacity using histograms (See Section 2.5.6).

Chapter 5 took the display further toward showing the objectives of the arrival management process. In this version, a secondary runway was introduced with the objective of minimising the use of that secondary runway. By equating deviation from an optimal arrival time against the use of a secondary runway, the *quality* of the trade-off between demand and capacity goals is made available to the operator.

7.3.3 Visualising uncertainty

The mathematical construct of an error distribution is technically simple to visualise since it is a 2D curve. The widths of these Probability Density Functions (PDFs) (up to 30 minutes) relative to typical landing intervals (2 minutes) cause the distributions to overlap (Section 6.3). The visual representation of the uncertainty of the arrival time itself is, therefore, difficult to read. Through the representation of demand as an area, the impact of uncertainty on the first objective of arrival management—spacing aircraft to fit on the runway—becomes much more salient. Participants in the experiments all understood that more spacing was required when the cumulative demand exceeded capacity (Chapter 4).

Looking back, Section 6.4 noted that the stacked demand curves do not fully constitute an expectation value. A demand curve that is higher than the capacity line does not indicate that a spacing conflict is certain. Similarly, a spacing conflict can still occur when demand does not appear to exceed capacity. However, the demand curves do provide an opportunity to calculate the likelihood of such a spacing conflict (Chapter 6).

Finally, the demand curves also provide a mathematical path to determine the working horizon objectively. Using the cumulative probabilities that a flight has arrived, the probability of a sequence swap may be calculated. This probability against time indicates the relationship between horizon and sequence stability (Section 6.4).

7.3.4 Evaluating an expert display for expert users

Neither the two experiments with a single time line (Chapter 4) nor the experiment with a display for multiple runways (Chapter 5) demonstrated an advantage of visualising uncertainty. Participants indicated difficulty in understanding the concept of uncertainty as presented on the display. A key problem here is the complexity of the work domain itself when uncertainty is introduced. In systems without uncertainty, the operator has to develop a mental model of the system to forecast how it *will* evolve given its current state. When uncertainty is introduced, the operator has to develop a different type of mental model: How it *could* evolve and how *likely* that is. In retrospect, the proposed risk indicators in Chapter 6 may well be the missing content from the work domain with uncertainty.

The resulting display is complex due to the fact the work domain itself is complex [14]. Since the work domain is complex, experiment participants have to become real experts in managing the system. While the prototyped interface may certainly not be the best way to visualise uncertainty in arrival management, any display would require novice users to receive considerably more training to fully benefit from novel interfaces.

Using the inspiration from EID did result in an indicator for higher-level performance: The quality of the trade-off in the goals of demand and capacity. In experiments, a non-significant improvement in performance was seen (Section 5.6). More importantly, the differences in performance between participants reduced as the worst performers benefitted most.

The need for sufficient data points introduced two deviations from the typical situation in which information on the uncertainty would be of value in arrival management (Section 3.7). First of all, most flights need close spacing when capacity is constrained. Creating buffers to accommodate uncertainty will lead to delays at the back of the queue. In current operations, the principle of First-Come-First-Served is often close to the 'practical optimum' at high traffic densities due to the lack of control space. Second, using faster-than-real-time simulation in our evaluations avoids the significant time delays between operator decisions and their consequences that would occur in a real two-hour arrival management process. This quick feedback between action and response allows for comparing the development of the situation against the operator's plan. In an actual operation, such a comparison would require the operator to retain a mental model for two hours.

7.3.5 Solving the right problem

With the display designed for a two-hour horizon and uncertainty as large as was found in Chapter 3, it could be argued that the task defined in the experiment is different from the task that is to be supported by the display. If uncertainty is as large as demonstrated, perhaps the task at longer horizons is more that of *flow control*. This process manages the number of aircraft arriving within a certain timespan but focuses less on individual flights.

Flow control in current operations is mainly limited to restricting departures. In European operations, flow control is executed by the NM on request of the ANSP. The ANSP sets a capacity for a certain airspace or airport. This capacity is set for a particular period and expressed in movements per hour. Subsequently, the NM predicts traffic load per sector and restricts departures when that load exceeds the available capacity. After departure, no further flight plan adjustments are applied in relation to this traffic load. Once the flight enters the destination ANSP's AOR, the arrival time will become a relevant control parameter again.

Two gaps exist between flow control as applied by the NM and arrival management:

- A gap in time horizon between the issuance of a departure slot and the time at which an aircraft is in the arrival management horizon,
- a gap in the controlled parameter between movements per hour for a given period and the exact landing time of a particular flight.

The proposed display bridges the gap in time. Furthermore, the aforementioned merge of the time line concept and the demand graph provides the visual representation of the transition from flow to individual flights.

However, the experiment only allows adjusting individual flights. Given the uncertainty of individual flights, adjusting flights one-by-one may not be the best type of action for an operator. A good example of a different strategy is the one recently (2017) introduced by the UK ANSP in which all inbound aircraft are requested to reduce speed by the same amount when overall demand is too high [15].

7.4 Recommendations

Uncertainty is unlikely to disappear altogether but can be estimated, modelled, and predicted using an empirical approach. The information on uncertainty may support strategies that allow for deviation at longer horizons while working towards more detailed solutions at short horizons. Furthermore, it allows dynamic planning horizons that adjust based on the actual operation. Such strategies may well be of interest in resolving the complexities of Trajectory Based Operation (TBO) in dense airspaces with shorter flight distances.

The uncertainty analysis in this thesis is based on a limited, and at the current time of writing (2022), perhaps even obsolete, dataset. It is worthwhile to explore more recent datasets and information sources. This research could address a more detailed classification of flights, for example, according to their A-CDM status. The research could also address prediction sources such as EUROCONTROL's EFD, downlinked information from aircraft, and perhaps even internal prediction capabilities developed within an ANSP.

The synthesis of the display and the data prompts a novel information item and a way to calculate that item: The stable sequence horizon. It is recommended that the proposed determination of the horizon is validated. Furthermore, the use of such a dynamic horizon in operations needs to be explored.

The experiment results showed no significant effects. However, many confounding factors remain. The most important factor is the operator's understanding of the display and the work domain. A redesigned experiment may well be more conclusive. The first recommendation for such an experiment is to provide much more training to each participant. A second recommendation is to use prediction errors with dynamics closer to actual operations to prevent the prediction error from becoming predictable.

Finally, it is worthwhile to consider the place of automation in arrival management at such horizons. As this dissertation shows, the complexity of the problem grows rapidly with longer horizons. It could be argued that solving the problem itself may well become too complicated for a human operator. In that case, automation would need to be developed to perform the arrival management task. In such a concept, knowledge of the risk of sequencing errors may help in dynamically adjusting the constraints of the automated planning approach.

If automation is to be applied, the role of the human operator transforms into that of a system *monitor*. If the system is too complex, a fully shared representation of the work domain cannot exist: The human operator will never be able to understand the lower level details of the problem for two-hour time horizons. However, a shared world model at a higher level could possibly be designed to allow the operator to determine the overall performance of the system and understand the full rationale of the automation. It is recommended that the proposed diagram is further studied as a route towards understanding the expected performance of arrival management at those two-hour horizons.

7.5 Conclusion

It is very unlikely that all disturbances in the planned trajectory of a flight will ever be predicted or removed. This especially holds for the part of the trajectory at the gate before departure. Hence, arrival management at longer horizons will always involve uncertainty. If arrival management is to be performed at a two-hour horizon, management of uncertainty will be an ever-present element in the decision making process.

By analysing historic prediction errors, it is possible to predict the uncertainty in a prediction of arrival time. Such a forecast allows online determination of the uncertainty per flight. The empirical method applies to predictions irrespective of the origin of the prediction and therefore works in a scenario where another ANSP provides the prediction. Since the parameters of the distribution are tabulated per class of prediction, the method is not computation-intensive and would work in an online environment.

Displaying the uncertainty on a time line diagram is complex. Experiments showed some performance improvement, but the results were not significant. The key issue in the experiment is building up participants' expertise in operating a novel display in a novel operation. The visualisation using the expectation of occupancy does provide a route toward indicating an effective arrival management horizon, given the uncertainty of the flights in the sequence.

The high uncertainty in the analysed dataset would limit the value of such a display at the proposed horizon in a scenario where the arrival management process is based on sequencing and merging flights. Reduction of uncertainty would probably improve the possibilities. However, a more likely route would be to adapt the process to better align with the gradual transition from flow control at longer horizons to arrival time control at smaller time horizons.

7.6 Bibliography

[1] A. Barff, B. Favennec, P. Conroy, *et al.*, "SESAR P05.06.04 - D28 - Preliminary OSED ed. 00.01.01," SESAR Consortium, Tech. Rep., 2012.

[2] J. Bronsvoort, G. McDonald, M. Paglione, C. Garcia-Avello, I. Bayraktutar, and C. M. Young, "Impact of missing longitudinal aircraft intent on descent trajectory prediction," in *30th IEEE/AIAA Digital Avionics Systems Conference*, Seattle, WA, USA, Oct. 2011.

[3] European Commission, *Commission Implementing Regulation (EU) 2021/116 of 1 February 2021 on the establishment of the Common Project One supporting the implementation of the European Air Traffic Management Master Plan*, 2021.

[4] N. Hasevoets and P. Conroy, "Arrival Manager - implementation guidelines and lessons learned," EUROCONTROL, Brussels, Belgium, Tech. Rep., 2010.

[5] "SESAR concept of operations at a glance ed 02.00.00," SESAR JU, Tech. Rep., 2011.

[6] JPDO, *NextGen integrated work plan: A functional outline*, 2008. [Online]. Available: `http://www.jpdo.gov` (visited on 03/01/2013).

[7] "Global Air Navigation Plan (GANP) 2016-2030," ICAO, Montréal, Tech. Rep., 2016. [Online]. Available: `www.icao.int`.

[8] J. Bronsvoort, G. McDonald, R. Porteous, and E. Gutt, "Study of aircraft derived temporal prediction accuracy using FANS," in *Proceedings of the 13th Air Transport Research Society (ATRS) World Conference 2009*, Abu Dhabi, UAE, 2009.

[9] R. A. Coppenbarger, "Climb trajectory prediction enhancement using airline flight-planning information," in *AIAA Guidance, Navigation, and Control Conference*, San Diego, CA, USA, 1999.

[10] G. Sölveling, "Stochastic programming methods for scheduling of airport runway operations under uncertainty," Dissertation, Georgia Institute of Technology, Jun. 2012.

[11] G. Tobaruela, P. Fransen, W. Schuster, W. Y. Ochieng, and A. Majumdar, "Air traffic predictability framework - Development, performance evaluation and application," *Journal of Air Transport Management*, vol. 39, pp. 48–58, 2014.

[12] N. L. Johnson, "Systems of frequency curves generated by methods of translation," *Biometrika*, vol. 36, no. 1/2, pp. 149–176, Jun. 1949.

[13] J. Rasmussen, "A taxonomy for analysis of cognitive work," in *IEEE Conference on Human Factors and Power Plants*, Monterey, CA, USA: IEEE, Jun. 1992.

[14] K. B. Bennett and J. M. Flach, *Display and interface design: Subtle science, exact art*. Boca Raton, FL, USA: Taylor and Francis, 2011.

[15] European Commission, *Extended Cross Border Arrivals Management optimising arrivals for London Heathrow Airport*. [Online]. Available: `https://ec.europa.eu/transport/modes/air/ses/ses-award-2016/projects/extended-cross-border-heathrow_en` (visited on 09/18/2019).

Acronyms

4DT	4-Dimensional Trajectory
ACC	Area Control
ADS-B	Automatic Dependent Surveillance Broadcast
AFTN	Aeronautical Fixed Telecommunication Network
AMAN	Arrival Manager
ANSP	Air Navigation Service Provider
AOR	Area of Responsibility
APP	Approach Control
ATC	Air Traffic Control
ATCO	Air Traffic Controller
ATD	Absolute Total Deviation
ATM	Air Traffic Management
AU	Airspace User
CDF	Cumulative Distribution Function
A-CDM	Airport Collaborative Decision Making
CDO	Continuous Descent Operations
CTAS-TMA	CTAS - Traffic Management Advisor
CTAS-EDA	CTAS - Efficient Descent Advisor
CTOT	Calculated Time of Take-off
DST	Decision Support Tool
DMAN	Departure Manager
EFD	ETFMS Flight Data
EID	Ecological Interface Design

ELDT	Estimated Landing Time
EOBT	Estimated Off-Block Time
ETA	Estimated Time of Arrival
ETFMS	Enhanced Tactical Flow Management System
ETE	Estimated Time Enroute
EU	European Union
FCFS	First-Come-First-Served
FIR	Flight Information Region
FMS	Flight Management System
FUM	Flight Update Message
HMI	Human-Machine Interface
IBP	Inbound Planner
IAF	Initial Approach Fix
ICAO	International Civil Aviation Organization
ISA	Instantaneous Self Assessment
KPI	Key Performance Indicator
LVNL	Air Traffic Control the Netherlands
MTCD	Medium Term Conflict Detection
NextGen	Next Generation Air Traffic Management System
NM	EUROCONTROL Network Manager
PDF	Probability Density Function
PTA	Planned Time of Arrival
PVD	Plan View Display
RBT	Reference Business Trajectory
RNAV	Area Navigation
RTA	Required Time of Arrival
SBT	Shared Business Trajectory
SESAR	Single European Sky ATM Research
STA	Scheduled Time of Arrival
STAR	Standard Terminal Arrival Route
SWIM	System Wide Information Management
TBO	Trajectory Based Operation

TMA	Terminal Maneuvering Area
TMC	Traffic Management Coordinator
TP	Trajectory Predictor
TOD	Top of Descent
TSD	Time-Space Diagram
TTA	Target Time of Arrival
TTD	Total Time Deviation
TTL	Time to Loose
TTL/G	Time to Loose/Gain
TTG	Time to Gain
UAC	Upper Area Control
URT	Unused Runway Time
WTC	Wake Turbulence Category

List of Figures

List of Tables

Acknowledgements

Here I am, typing the last legs of a trajectory that took off in the summer of 2011 and had some uncertainty in its ETA. And when you have spent eleven years on a dissertation, you may need to thank a person or two. There is no way I am going to be sure that I mention everyone, though. So I start with a general thank you to everyone I met during this work. So many—as of today, including my five-year-old son—have asked when my dissertation would be finished. Whether in jest, desperation, or curiosity, it motivated me to persist and finish the bloody thing, and I thank them for it.

However, here are some people I want to mention in particular. Let me start with Max, the man who got me into doing this and helped me all the way to the end. Most reviewed drafts carry his handwriting and progress or no progress; he was always available for a conversation. And when I mention Max, I cannot forgo my other promotor, René. No problem is too complex for this brilliant mind, yet no question too simple to help out with. Beyond these two, I thank all current and former colleagues at Control and Simulation for coffee breaks, inspiration, hopelessly bad basketball, adventures in zero-G and willingness to be subject in my experiments. Rolf, it has been a long road, but you'll get there as well.

The KDC sponsored this thesis, and with that comes access to arrival management in real life. I thank Evert for being a sponsor with limitless patience. Although, I suspect that working in the innovation of ATM hones such a skill to perfection. I thank Dennis for starting me up in the right direction and giving access to knowledge and experts at the NLR, the team at LVNL for allowing me to work on the real-world system and Ferdinand, with whom I keep fantasising about how we would build the ATM systems if we would be given the chance.

My transition to To70 allowed me to keep working on the thesis and in ATM. The enduring support in means and motivation by Ruud and Kjeld allowed me never to abandon the dissertation and work on it whenever the situation allowed. No longer will my colleagues joke about its completion. But even when joking, they showed their support. A special mention to Ian for helping me out with the statistics of the probability of a sequence switch, which led to a whole new idea also explored in this work. Further shout-outs to Adrian and Matthijs for their reviewing work.

And I thank my friends, of whom I count myself blessed to have so many. Some of you may have introduced some unforeseen delay, particularly those who sat on the bench at the Vredespaleis and convinced me to take a different—and in hindsight, very consequential—side tour. You provided me with lots of reasons not to work on this dissertation, but still your support got me through. This is also where I would be remiss if I did not mention the relation between thirds of perpendicular angles, the importance of minutes, and world domination, if only because my successors now will have to as well.

I thank my family, who were always supportive, sometimes joking, sometimes by taking care of the kids and sometimes by taking care of me. And this book is also for my grandparents. I know that they are proud. This is their legacy as well.

And finally, Margot. Life is a journey, and somewhere on that road the two of us found each other. We've seen our share of good times and our share of bad times. But in both, you managed to keep my motivation for this adventure going and helped wherever you could. Many uncertainties in our trajectories have been reduced, let's see what is beyond the horizon.

Biography

Maarten Tielrooij was born on 30 April 1981 in Groningen in The Netherlands and grew up in Uithuizen. From 1993 until 1999 he attended Het Hogeland College in Warffum where he obtained his Atheneum Diploma.

In 1999 he enrolled at the Faculty of Aerospace Engineering at the Delft University of Technology. It was only during his internship at Airservices Australia—the Australian ANSP—that the subject of air traffic control piqued his interest. In 2008, this interest resulted in a Master's thesis titled "Assisting ATC in supervising continuous descent approach procedures".

Following his graduation, he moved to the United Kingdom to work for NATS En-Route Ltd.—the UK's ANSP—as a system engineer on trajectory prediction algorithms. Part of his work included SESAR projects on improved prediction accuracy using airline-provided mass information and development of Medium-Term Conflict Detection for advanced RNP manoeuvres.

In 2011, Maarten returned to the Faculty of Aerospace Engineering at Delft University of Technology to pursue a Doctor of Philosophy in Aerospace Engineering. As part of the research, he worked on LVNL's new arrival manager, directly applying the knowledge gained into operation.

Since 2016, he works as senior aviation consultant at To70 Aviation B.V. in The Hague. In this role, he advises airports, ANSPs, airlines, and government entities using his knowledge of aviation and air traffic management in particular. As a consultant, he has worked on the final specification, verification, and validation of LVNL's arrival manager, which is now operational. Maarten regularly works on projects aimed at extending the arrival management horizon and the implementation of Trajectory Based Operations.

List of publications

[1] M. Tielrooij, "Assisting ATC in supervising continuous descent approach procedures," MSc. Thesis, Delft University of Technology, The Netherlands, 2008.

[2] M. Tielrooij, A. C. in 't Veld, M. Mulder, and M. M. van Paassen, "Assisting ATC in metering, sequencing and merging arrival streams in three degree decelerating approaches," in *Proceedings of the 27th European Annual Conference on Human Decision-Making and Manual Control*, Delft, The Netherlands, 2008.

[3] M. Tielrooij, A. C. in 't Veld, M. M. van Paassen, and M. Mulder, "Development of a time-space diagram to assist ATC in monitoring continuous descent approaches," in *Air Traffic Control*, M. Mulder, Ed., Rijeka, Croatia: InTech, 2010, pp. 135–147. [Online]. Available: `http://www.intechopen.com/books/air-traffic-control/development-of-a-time-space-diagram-to-assist-air-traffic-controllers-in-monitoring-continuous-desce`.

[4] A. M. P. de Leege, M. Tielrooij, A. van der Eijk, and M. Mulder, "A time-space diagram based controller support tool for continuous descent operations," in *The International Conference on Human-Computer Interaction in Aerospace*, Brussels, Belgium, Sep. 2012.

[5] M. Tielrooij, C. Borst, D. Nieuwenhuisen, and M. Mulder, "Supporting arrival management decisions by visualising uncertainty," in *Proceedings of the SESAR Innovation Days*, D. Schaefer, Ed., Stockholm, Sweden, 2013.

[6] M. Tielrooij, C. Borst, M. M. van Paassen, and M. Mulder, "Supporting arrival management decisions by visualising uncertainty," in *17th International Symposium on Aviation Psychology*, Dayton, OH, USA, 2013, pp. 534–539.

[7] J. De Wit, M. Tielrooij, C. Borst, M. M. Van Paassen, and M. Mulder, "Supporting runway planning by visualizing capacity balances of arriving aircraft streams," in *Conference Proceedings - IEEE International Conference on Sys-*

tems, Man and Cybernetics, San Diego, CA, USA: IEEE, Oct. 2014, pp. 2989–2994.

[8] M. Tielrooij, C. Borst, M. M. van Paassen, and M. Mulder, "Predicting arrival time uncertainty from actual flight information," in *11th USA/Europe Air Traffic Management Research and Development Seminar*, FAA/EUROCONTROL, Ed., Lissabon, Portugal, Jun. 2015, pp. 557–586.