Please fasten your seatbelts

Data analysis towards the origin of departure delay
of KLM Cityhopper at Schiphol Airport

Master thesis Industrial Engineering & Management

Author:

Matthijs Kieskamp

UNIVERSITEIT TWENTE.

KLM

Please fasten your seat belts

Data analysis towards the origin of departure delay of KLM Cityhopper at Schiphol Airport

By

M. H. Kieskamp

Amsterdam, 3 December 2015

Master programme: Industrial Engineering & Management
Specialisation track: Production & Logistics Management
Institute: University of Twente
Faculty: Behavioural, Management and Social Sciences (BMS)
Company: KLM Royal Dutch Airlines
Department: ITCDOR

Committee:

Dr. Ir. L.L.M. van der Wegen   University of Twente
Dr. Ir. J.M.J. Schutten   University of Twente
A. Fiaz, MSc   KLM ITCDOR
Drs. J.A.C. van der Veer   KLM Cityhopper

UNIVERSITEIT TWENTE.   KLM

UNIVERSITEIT TWENTE.

Preface

Before you lies my graduation thesis. With the submission and defence of this thesis, I hope to end my years as a student and join the KLM as a full time employee. When I started by bachelor at the University of Twente in September 2008, I did not have the slightest idea on what a wonderful experience the next 7 years would become.

I chose my bachelor because I liked the combination of IT and management, and took the extensive mathematics for granted. However, after a few years, I started to realise that applying this mathematics into operations research models is something I really enjoy. Therefore, when the time came and I had to chose my masters specialisation, there was no doubt that it should be the Production and Logistics Management track. During my master, not in the least during my exchange semester at the Technical University of Munich, I discovered that the most interesting cases are the large and complex logistical puzzles. Therefore, I knew that I wanted to graduate at a large company, preferably with high tech production lines. Well, there are no production lines at KLM, but it is big and high tech, and I like it. The immense scale and precision keeps surprising me every day. Every day, KLM transports over 50,000 passengers. Every day! The past 7 months have been a continuous journey, with new discoveries every day.

I take this opportunity to thank some people. First and foremost, I thank Anne-Jasmijn for her support and understanding the past few months. She knew when to apply some brakes, or when she had to let me go. Also, she could listen to my continuous rambling about aircraft without any complaints (something that will not stop in the foreseeable future...) Next, I thank my parents for giving me the change to attend university. Without their mental, physical, and financial support, I would not have been here, Thanks!

Of course, I thank my supervisors, both at KLM and the university. Adnan, Jan, thank you for the daily support you provided and not growing tired of the endless stream of questions (or, at least not showing it...). Leo, Marco, thank you for the freedom you gave me in defining and executing this research, while at the same time offering valuable coaching and guidance. I really feel that each of our meetings (although not very regular) were very useful and provided me with ample input to continue.

Finally, I thank all colleagues at both KLM Cityhopper and KLM Decision Support for the warm welcome I received and the continuous support. I really feel I could ask anything to anyone at any time. Arjan and Joris, thank you for your input and keeping me sharp!

Amsterdam, December 2015

UNIVERSITEIT TWENTE.

Management Summary

Introduction

KLM Cityhopper (KLC) is in a challenging position. Increasing competition by low lost carriers makes it necessary for the traditional airliners to provide additional value for their customers to keep their premium position in the market. One of these added values is departing and arriving on time. This way, passengers can have the peace of mind that they will reach their transfers in time, or that they will be home in time for dinner. Therefore, one of the seven key performance drivers of KLM Cityhopper is punctuality.

However, over the past year, departure punctuality has seen a decline and KLC management wants to gain insight into why this is happening and what they can do to prevent departure delay in the future. However, the problem is that the turnaround process (the process of servicing an incoming aircraft and preparing it for departure again) is a very complex process with many activities, dependencies and parties involved. Thus, numerous places where delay can occur. This makes it hard for KLC to understand why departure delay deteriorates and what they can do to increase punctuality again.

Goal and approach

This report tries to unravel this complex process and find improvement suggestions for KLC management. To this end, we have the following research question:

Where can KLM Cityhopper best focus its efforts to improve departure punctuality?

To answer this question, we take a four-step approach. First, we gather data on the turnaround process of KLC at Schiphol Airport. Second, we clean our data to make sure our data is reliable and complete. Next, we apply a machine learning technique called decision tree learning on our data to find correlations between the different variables in our data and departure delay. The fourth and final step of this research is a simulation study. Using the results of our third step and input by KLC management, we test various potential improvements to see whether they indeed have a significant effect on departure punctuality.

Results

In our first step (Chapter 2), we learn about the turnaround process of KLC at Schiphol Airport. There are 15 turnaround types, determined by aircraft type, planned ground time, aircraft location, and gate location, each with different performance requirements. A complicating factor in planning, scheduling, and performance monitoring of the turnaround process is that all different activities that take place during a turnaround are performed by a different service department. Nonetheless, KLC has a thorough overview of how the turnaround process should happen and has made agreements (both qualitative and quantitative) with each of these departments towards this goal. However, the current way of monitoring performance (delay codes) can only give a general impression and is by no means an extensive performance monitoring tool. This results in KLC employees monitoring random turnarounds by hand, with a stop watch and performance sheet. Our data includes 70,000 turnarounds between 01-01-2014 and 01-06-2015. Removing unreliable or missing data (Section 3.4) results in a final dataset of 47,000 turnarounds.

Our data analysis (Chapter 4) results in multiple variables that correlate highly with departure delay (Table 1). Most of these variables concern activities that take place at the end of the turnaround process. A possible explanation is that arrival oriented activities still have some slack time whereas departure oriented activities do not have such a slack, making any disturbances result in departure delay.

Variable	Description
Local ATC lateness	Time between aircraft ready and actual departure
Passenger door closed & cargo door closed lateness	Time at which the various doors were closed, compared to their scheduled closure
Ready to depart lateness	Time between all doors closed and ready to depart notification
Arrival punctuality	Time the inbound flight was delayed
PaxScan Dminus	Time between last boarding pass scan and scheduled departure
Pushback start Dminus	Time between the start of the pushback task and scheduled departure
Load factor	Percentage of seats occupied

Table 1 - Variables correlating with departure delay

This is confirmed by the results of our simulation study (Chapter 5). Here, we see that changes to the final stages of the turnaround process result in an increased departure punctuality. The best improvements are measured when we introduce a ground time buffer between the departure oriented tasks and scheduled departure. This way, any disturbances during the final stages of the turnaround are absorbed, also the most important influence on departure punctuality, Local ATC lateness. The turnaround process is designed to receive instant departure clearance when requested, thus any ATC lateness will result in a delayed departure. However, especially during peak hours, over 5 minutes of waiting time can be present. Next to the scenarios with the added ground time buffer, scenarios where we reduce or even eliminate this lateness result in a greatly improved departure punctuality.

Starting the boarding pass scanning process earlier is the next best improvement. By starting earlier, passengers arrive at the aircraft earlier, allowing them more time to find their seat, stow their luggage, and fasten their seatbelt. At the same time, combining various turnaround types into one results (besides a greatly reduced complexity) in a similar improvement if KLC choses to use the current D-gate turnaround schedule as template. However, we must note that in our simulation, we treated each turnaround as a single entity without interdependencies or capacity restrictions. Also the availability of gate agents and busses should be checked. However, if there is enough capacity available, we highly suggest a pilot to gather real life results.

	Scenario	Results with regard to departure punctuality
1	Original	-
2	Increased ground time for QTAs	Only a slight improvement when used as buffer between arrival and departure oriented tasks, major improvement if used as a buffer right before STD.
3	Altered arrival punctuality	Slight improvement when arriving 5 minutes earlier, Serious deterioration when arriving 5 minutes later
4	Reduce 15 turnaround types to 3, QTA, standard E90 and standard F70	Depending on chosen new schedule, serious improvement at Platform and semi contact turnarounds
5	Altered boarding pass scanning process	Improvements at all turnaround types for both earlier start and shorter duration. Biggest gain at platform and semi contact turnarounds.
6	Local ATC improvement	Serious improvement at all turnaround types when reducing lateness to zero minutes.

Table 2 - Simulation results

UNIVERSITEIT TWENTE.

KLM

Table of Contents

UNIVERSITEIT TWENTE.

UNIVERSITEIT TWENTE.

KLM

Chapter 1: Introduction

This chapter starts in Section 1.1 with a general introduction of KLM, where this research is performed. It continues in Section 1.2 with an introduction to the problem that was the reason for this research, resulting in the research scope and – objective in Section 1.3. It finishes in Section 1.4 with providing the research questions that will guide the rest of this report.

1.1   KLM – Company description

KLM Royal Dutch Airlines is the oldest airline in the world flying under its original name. Since its founding in 1919, the airline has served regular flights from its main airport Schiphol Amsterdam (AMS) all over the world. In 2004, the company merged with Air France into the holding Air France-KLM. Excluding daughter companies, KLM has a fleet of 138 aircraft, consisting of Airbus 330 and Boeing 737, 747, and 777 aircraft types. In 2013, KLM transported 26.5 million passengers and 770,000 tons of cargo with a workforce of over 30,000 FTE. Daughter companies of KLM include KLM Asia, KLM Cityhopper, KLM Cargo, Martinair, and Transavia.

1.1.1   KLM Cityhopper

KLM Cityhopper (KLC) is a 100% daughter company of KLM Royal Dutch Airlines. It generates more than half of the turnover of the KLM group in Europe. It does this by operating 47 aircraft on 287 daily short to medium haul flights to 53 European destinations from its hub airport, Schiphol Amsterdam (AMS). As can be seen in Figure 1, these destinations are mostly located in Western and Central Europe. By providing a high frequency of flights to major business hubs such as Dusseldorf, Munich, and London, KLC positions itself towards business users and provides a transfer service between its airports and Schiphol, to enable intercontinental (ICA) connections with the KLM Group. On average, KLC transports over 18,000 passengers daily. Its fleet consists of 19 Fokker 70 and 28 Embraer 190 aircraft with a maximum of 80 and 100 passengers respectively. In March 2015 KLC announced to modernize its fleet by replacing the aging Fokker aircraft with modern and more fuel efficient Embraer 175 and 190 aircraft, thus planning to have its entire fleet consist of Embraer aircraft by the end of 2018.

Figure 1 - Current KLC destinations

1.2   Problem description

Departure punctuality is an important performance indicator for airliners. It is often expressed as the percentage of flights departing on or before its scheduled time of departure. Delay at departure indicates either burning more fuel to try to catch up or arriving with delay at the destination. The first means higher costs and almost certainly makes the flight unprofitable, the latter means delaying connecting flights of passengers, crew and aircraft, also resulting in higher costs. Research suggests (Cook, Tanner, & Lawes, 2012) that punctuality is considered a 'basic factor', meaning it does not generate satisfaction when fulfilled, but generates dissatisfaction when it is not. This departure punctuality has been deteriorating the past few years at KLC. In Figure 2, we clearly see a downward trend in the departure punctuality (D0 performance) with July 2014 as low point (scale removed due to confidentiality).
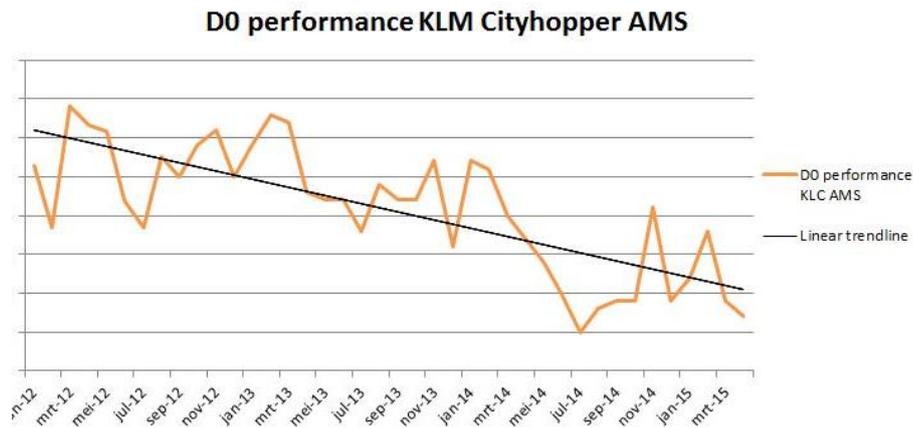
Figure 2 – D0 performance KLC departing from AMS

The aircraft of KLC perform up to seven or eight flights each day and a disturbance early on the day may influence the schedule of the entire day. This makes turnarounds (the process of preparing an aircraft for departure right after arrival) critical for the performance of KLC as it is believed by management that delay often originates here. During a turnaround, many activities have to be performed in order to get an aircraft airborne again. A few examples include: fuelling, cleaning of the cabin, replacing the catering, emptying the lavatory, (un)loading of luggage and passengers, placement of safety cones around the aircraft, transport of handicapped passengers, and pushback of the aircraft before departure. Appendix B contains a full list of all events and activities that take place during a turnaround. All of these activities are places where delay can occur. A complicating factor is that each activity is performed by a different department or contractor, each with its own planning, capacity, and priorities. Currently, just like at many other airliners (Wu C.-L. , 2008), the turnaround process of KLC has been well documented internally but actual performance is a 'black box' as information regarding the performance of these activities (if there is any) is only available at the contractors themselves without a general information / knowledge sharing platform. This makes it hard for KLC to understand where the deterioration of departure punctuality originates and thus how it can improve its departure punctuality.

Delay identification tries to pinpoint the origin of delay. If an airline gains insight in where delays originated, it can work on taking away the cause, thus improving future performance. To this end, the International Air Transport Association (IATA) has developed a standard delay coding system to store causes of delay. In this coding system, a flight gets a double digit code assigned when it has been delayed. This delay code corresponds with a cause of delay (Wu & Truong, 2014). Like most airlines, KLC has expanded and tailored these delay codes it to its own needs, resulting in 87 delay codes available for operators to assign to delayed flights. Each delay code is matched with a responsible department to enable easy performance review. Appendix C contains a full list of the delay codes being used by KLC.

However, the current delay identification method is neither exhaustive nor reliable (see also Sections 3.1 and 4.3). Therefore, a project team was assigned to gather historical information regarding the performance of turnarounds from the different contractors. This project, dubbed AP2K (Analysis Process Performance KLC), was initiated to provide insight into where delay originates. Based on the Business Intelligence tool Spotfire, a data visualisation tool with statistical and mathematical capabilities, an application is being developed where all available data from the past two years is gathered and all future performances will be stored as well, creating a possibility to run daily performance analysis. Users will be able to zoom in onto specific turnarounds, get an average performance of all turnarounds in one day / week / year or

2

look at specific activities. The data from over 117,000 turnarounds has been added to the system so far. The tool is built to calculate the origin of delay for each turnaround by determining contributed and propagated delay, meaning that activities that delay other activities will account for the total delay they caused, see Section 3.2.1. However, while providing already some good insights, the tool currently does not provide insight into trends or specific situations which might cause delay and cannot do anything 'smart' with the data. However, a lot of information can be gathered by smart analysis, which KLC can use to improve its long-term performance.

1.3   Research scope & objective

This research focuses on providing insight where KLC might improve its departure punctuality (D0 performance). Because this research is part of the AP2K project and to make sure the scope does not become too large for one researcher in five months, the input for this research is limited to data gathered in the AP2K project and data gathered from interviewing internal stakeholders. If data do not match these criteria, it is considered out of scope. Currently, only data from turnarounds performed at Schiphol is available. Therefore, the scope of this research excludes outstation (non-Schiphol) performance. Important to note is that this research focusses on finding correlations between circumstances and departure delay. It is not meant as a root-cause analysis into the actual origin of delay. The objective of this research consists of two intertwined parts, namely to provide KLC with insight into the current performance of the different turnaround activities at Schiphol, and insight in where KLC can best focus its efforts to improve its D0 performance.

1.4   Research questions & approach

The leading question during this research is as follows:

> *Where can KLM Cityhopper best focus its efforts to improve departure punctuality?*

To answer this question, we perform a threefold research. First, we gather data regarding historical performance and look how turnarounds are performed at KLC. We then analyse the current performance based on this data to look for possible causes of delay. Finally, we simulate the critical path of a turnaround to see where process improvement would be most effective. These three steps each have their own sub research question that, along with a short description and plan of approach on how each of them is answered, are:

1) What is the required performance level of the turnaround process of KLC at Schiphol airport?
   a) What activities are part of the turnaround process?
   b) What key performance indicators are currently used to evaluate performance?
   c) What are the currently required performance levels on these KPIs?

To answer this question, we analyse the turnaround process of KLC using internal documents such as the OCP-ZY20 and Service Level Agreements with suppliers, added with interviews with key stakeholders. Based on this information we present an overview of activities that are performed during turnarounds at Schiphol, their key performance indicators (KPIs) and current required performance on these KPIs. We discuss Question 1 in Chapter 2 of this report.

2) What data is required to assess the actual performance of the turnaround process?
   a) How is delay identification currently performed?
   b) Which circumstances or activities are suspected or known to have a serious impact on D0 performance?

      c)   What data is available regarding the historical performance of activities during a turnaround?

      d)   Which data can be used for the performance assessment?

In this step we look at which data we need to perform our assessment and which data is actually available. We look at the current delay identification methods and what circumstances are known or suspected (by either KLC or literature) to have an impact on D0 performance, again using internal documents and interviewing key stakeholders supplemented with literature research. Based on this, we try to identify additional KPIs alongside the ones currently in use at KLC to fully assess D0 performance. Next, we look at the data gathered by the AP2K project, assessing both the quality of the data and what it actually represents. Using the answers to Questions 1) and 2a) unto 2d), data completeness and data quality, we define a dataset on which the performance assessment of Question 3) can take place.

As input for answering Question 2) we use the current delay identification method for Question 2a), literature and interviews with different stakeholders: two Duty Managers Ground Services (DMGS) and the Manager Operational Processes & Planning for Question 2b) and the data gathered by the AP2K project for Question 2c). We discuss Question 2 in Chapter 3 of this report.

    3)   What is the actual performance of the turnaround process?
      a)   What is the current general performance of each activity?
      b)   Which circumstances or activities actually have an impact on D0 performance?
      c)   Is the current delay identification method reliable?

Based on the data set that results from Question 2d), we perform the actual performance assessment. We use the OCP-ZY20 document as a baseline performance of the entire process and its different activities and compare this to the actual performance in Section 4.1.1. After that, we analyse our dataset to see which measurements show a correlation with departure delay and thus might be a cause of delay. We perform this analysis using decision tree learning, a data mining technique (see Section 4.2.1). Next, we test whether flights currently receive the right delay code, by checking whether certain time measurements correspond with the assigned delay code for that flight. We discuss Question 3 in Chapter 4 of this report.

    4)   Which improvements would contribute to a better D0 performance the most?
      a)   How can we test the expected performance of various improvement efforts?
      b)   What improvements show the highest expected improvement?

The final step in this research is to test how the D0 performance would improve if certain changes were implemented. The changes we test are based on the findings of Chapter 4 and requests of KLC senior management. To this end, we construct a simulation model of the turnaround process, using historical data as input. Based on the results of these simulations, we conclude with advising KLC where it should focus its efforts to improve D0 performance effectively. We discuss Question 4 in Chapter 5 of this report.

Chapter 6 of this report starts with answering Questions 1 through 4. In the second section of this chapter, we answer our main research question: "Where can KLM Cityhopper best focus its efforts to improve departure punctuality?", after which we discuss the practical implications for KLC. Finally, we discuss the limitations of our research and some recommendations where further research could prove its worth.

A short hint on navigation through this report: To improve coherence and readability of the report, we place each sub question as sub title of the section where it is answered. For example, the sub question of Question 1 'What activities are part of the turnaround process?' is answered in Section 2.1. We therefore place this question as sub title underneath the section title.

UNIVERSITEIT TWENTE.

Chapter 2: Turnaround Processes at KLM Cityhopper

In this chapter, we look at the first research question: What is the required performance level of the turnaround process of KLC at Schiphol airport? We describe the current turnaround processes at KLC and provide an overview of the different activities involved in preparing an aircraft for take-off in Section 2.1. We then look at how KLC currently measures performance (Section 2.2) and what service level agreements have been agreed upon with service providers (Section 2.3).

2.1   The Turnaround Process
What activities are part of the turnaround process?

As mentioned in Chapter 1, the turnaround process is the process where an arriving aircraft is serviced and prepared for a new flight. Dependent on the time of day, maintenance requirements, aircraft type, flight and airport, this process can take up anywhere between half an hour to multiple days. In case of KLC, the default turnaround processes take place in 35 to 50 minutes due to the short- to medium haul flights that it operates, combined with the high frequency of flights. Aircraft often perform over seven flights on a single day, making reliable turnarounds crucial for KLC to remain profitable.

Important to note is what we define as the actual turnaround process. Regular airline definitions define (Wu & Craves, 2004) the turnaround process as the time between 'on blocks' and 'off blocks', the time the aircraft is parked at the gate or on the platform. All time spent taxiing from and to the runway, waiting for a free spot at the gate or de-icing before departure is not included in the turnaround time. All other activities are thus part of the official turnaround process, including waiting at the gate for clearance from air traffic control (ATC) to start taxiing for take-off.

During planning, total turnaround time is based on two parts, the minimum ground time and schedule buffer time (Kohla, Larsen, Larsen, Ross, & Tiourine, 2007). The minimum ground time is the time needed to perform all activities needed before take-off. Schedule buffer is additional idle time, either to buffer for uncertainty in the schedule or simply because the aircraft is not needed sooner. During the minimum ground time, a large number of activities have to be performed. A full list of activities and events performed by KLC during a turnaround at Schiphol can be found in Appendix B, a schematic simplification of the process can be found in Figure 3. The most important activities include on- and offloading of passengers (Pax), on- and offloading of luggage (Bax) and cargo, replenishing the on-board water supply, emptying the lavatory, cleaning the cabin, performing a maintenance check, replenishing the catering, refuelling the aircraft, delivering the load sheet (list of all items and persons aboard the aircraft), and pushing back the aircraft from its position.

Activities are either planned as arrival oriented or departure oriented. This means they are planned to be executed with either a certain number of minutes after arrival or a certain number of minutes before departure. A clear example is the unloading and loading of passengers. The first has to be performed right after arrival while the latter has to be performed right before departure. In quick turnarounds, arrival and departure oriented timespans can be adjacent or, when dependencies allow it, overlap. Given the focus of this research on departure punctuality, we are only interested in arrival oriented activities when they may interfere with departure oriented activities and thus departure punctuality. In practice, this means that we will ignore arrival activities at the end of the day and those that are part of longer turnarounds.

Of course, there are also other dependencies among the different activities. There are obvious ones, such as loading of new luggage is not possible when the old luggage is still on board, regulations that dictate that, for example, the water supply cannot be replenished at the same time the lavatory is emptied, and practical limitations that the cabin cannot be cleaned while on- or offloading passengers. Dependencies between activities can also be found in Appendix B.
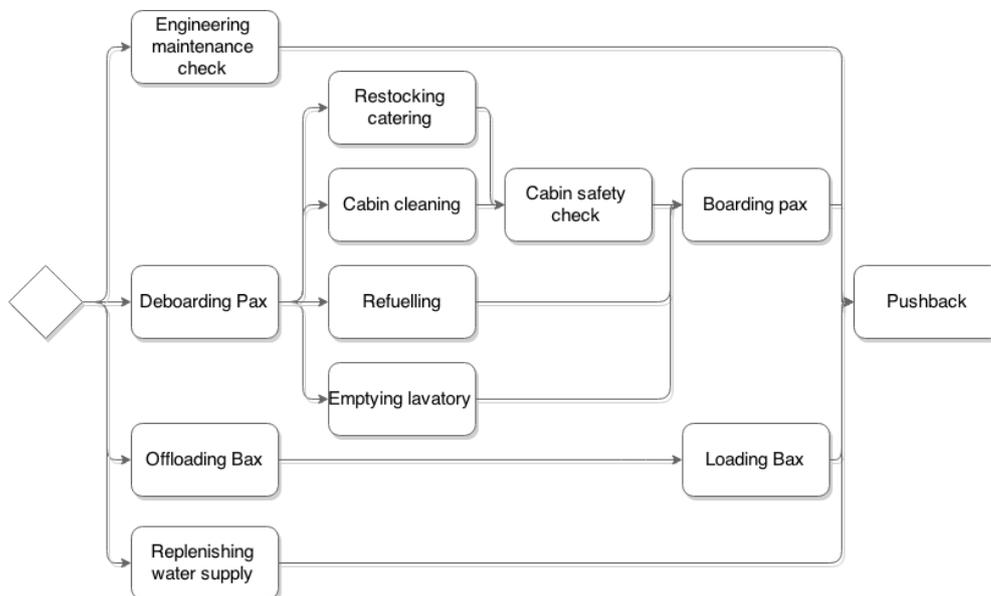
Figure 3 – Simplified turnaround process with precedence relations

Another complicating factor and important to note is that the servicing during a turnaround is not performed by KLC itself but instead purchased from local ground handling service providers. In case of Schiphol Airport, this third party is KLM Ground Services (KLM GS). KLM GS in turn divides its activities over multiple departments: The scheduling is performed by KLM Tactical Planning (KLM ST) and operational planning of the different activities is performed by Hub Control Centre (HCC), actual execution of the tasks is done by Aircraft Services (KLM AS). However, KLM AS uses third parties to perform some activities, for example, the catering is replenished by KLM Catering Service (KCS) and cleaning is performed by the company Klüh. On an operational level, KLC monitors performance of the different contractors and guards whether the different subcontractors keep to their SLAs. However, due to the complex organisational structure, pinpointing exact delay causes can be cumbersome if no Duty Manager Ground Services (DMGS) of KLC is present to assess performance.

2.2 Current performance indicators

What key performance indicators (KPIs) are currently used at KLC to evaluate turnaround performance?

Turnaround performance is measured and assessed with various sets of KPIs, dependent on the management level and organisational area. In this section, we use data that we have gathered from 8 reports used at KLC, being:

- The weekly report regarding ground process performance as created by the Manager Operational Processes & Planning to be used by the Director Ground Services & Cargo Operations KLC. (appendix D)
- Monthly non-performance overview Ground Services (ZY) on all stations, also created by the Manager Operational Processes & Planning to be used by the Director Ground Services & Cargo Operations KLC. (Appendix E)

6

- Monthly non-performance comparison of Ground Services at Schiphol, non-performance compared to same month last year. Again created by the Manager Operational Processes & Planning to be used by the Director Ground Services & Cargo Operations KLC. (Appendix F)
- The weekly report regarding ground process performance in Amsterdam, supplemented with specific service provider activity overviews, created by the various Duty Managers Ground Services to be used by widely in the organisation, but also by the Director Ground Services & Cargo Operations KLC. (Appendix G)
- The monthly summary of performance conformance during turnarounds monitored by the Duty Managers Ground Services, again to be used by the Director Ground Services & Cargo Operations KLC. (Appendix H)
- Performance overview KLC Flight Operations (ZV, the department responsible for cockpit crew and flight performance) used by (among others) Director Flight Operations. (Appendix I)
- Weekly report regarding operational performance created by Operations Control (ZZ, the department responsible for managing daily operations network wide) to be used in weekly management team (MT) meeting of KLC. (Appendix J)
- Monthly report from KLC management to general KLM management regarding KLC's overall performance.

Important to note is that there are more reports used at KLC that assess performance but unfortunately there are too many to all be taken into account. Therefore, we made a selection based on highest importance for KLC, relevance for this research, and the least amount of overlap between reports.

	Driver	Description (KLC internal)
1	Safety	Never compromise on Safety
2	Compliance	To obey the law
3	Completion	Percentage of flights accomplished according to schedule
4	Punctuality	Flights departing / arriving on time
5	Customer satisfaction	Passengers that experience "value-for-money" resulting in willingness to fly with us again
6	Employee satisfaction	Professionals who are engaged, fit for their jobs and proud to work at KLM Cityhopper
7	Cost per unit	Lowest total cost per flight, passenger and available seat kilometre

Table 3- Performance drivers KLC

The large emphasis on reports from Ground Services can be explained from the primary performance drivers of KLC and the activities of the different departments. KLC has seven major drivers on which it bases its daily operations (Table 3). As we see, departure punctuality is only part of the fourth important driver, after safety, compliance and completion. Each department has a different influence on these drivers, with Ground Services being the department with the biggest influence on departure punctuality as they handle the turnaround. Nearly all activities of Ground Services influence departure punctuality, directly or indirectly. Operations Control can influence departure punctuality by cancelling a flight or performing a registration change. In a similar fashion, Flight Operations only has influence on departure punctuality through arrival punctuality and the timely arrival of the cockpit crew before departure.

UNIVERSITEIT TWENTE.

2.2.1 Performance indicators

As expected, departure punctuality is an important performance indicator on many of the examined reports. It is expressed as D0 and D15 performance, where the number after the D indicates the number of minutes after scheduled departure time. For example, a D0 performance of 40% indicates that 40% of all flights left on or before their scheduled time, whereas a D15 performance of 85% indicates that 85% of all flights left on of before 15 minutes after scheduled departure. KLC has set targets to have a certain percentage of flights within both categories, actual performance is compared to these targets.

The second important conclusion drawn from these reports is that delay cause identification is currently performed using delay codes. A delay code and its corresponding delay length are assigned to a flight to indicate the source of a delay and the amount of delay this has caused. By matching delay codes to their corresponding responsible departments, KLC monitors its internal performance and these departments in turn use them to measure performance within their own department. However, as we discuss in Section 3.1, the assignment of the delay codes is done by hand and thus subjective.

Therefore, in addition to the delay coded performance assessment, a Duty Manager Ground Services (DMGS) from KLC samples the performance of 50 turnarounds each month by hand. In these samples, various soft- and hard performance criteria are assessed to be either satisfactory or unsatisfactory, accompanied by an optional remark. For example, the DMGS looks at whether the various service agents finish their job in time, but also whether they park their vehicle at the right location or check the platform for foreign object debris (FOD). The full checklist can be found in Appendix K.

These 50 samples are aggregated and this data is used as input in the monthly performance meeting with the various subcontractors. Important to note is that the on-time performance of activities is only measured by completion time of the activity, ignoring starting time and activity duration; only a very late start is mentioned in the remarks as an explanation of delay. This is consistent with the agreed service levels with the service providers that we discuss in Section 2.3.

Flight Operations has created a periodic dashboard where they measure their own contribution to KLC's departure punctuality. In this dashboard, the sum of all delays (presumably) caused by Flight Operations is divided into four types of delay, based on thirteen possibly assigned delay codes. Flight Operations identifies four possible types of delay: 'Local Air Traffic Control (ATC)', 'en-route ATC', 'flight plan & fuel' and 'crew'. The precise meaning of these delays can be found in Section 3.1. Note that only the last two are under direct control of Flight Operations as ATC is an external party to which Flight Operations has to obey.

Operations Control uses a dashboard with various performance indicators, grouped to the seven important performance drivers of KLC. Regarding safety, they measure the number and severity of incidents. For completion, they measure the percentage of flights that was not cancelled within 72 hours before scheduled departure. As mentioned, this metric is more important to KLC than their punctuality. This means that, depending on circumstances, Operations Control will in general prefer to delay a flight rather than cancel it.

For most performance indicators, specific targets are set each period based on historical performance and perceived potential improvements. The performance is compared to both this target and sometimes the same period the year before to allow fair comparison. The comparison is not based on the period right before (e.g. March to February) because of seasonal effects.

UNIVERSITEIT TWENTE.

In the monthly KLC performance summary as presented to the mother company KLM, performance on all seven performance drivers is discussed together with financial summaries, project statuses, business risks and other general management information. Because of this wide scope, regarding punctuality only the major performance indicators are presented, namely D0, D15, A0 and A15.

A general overview of all KPIs that are relevant for this research can be found in Table 5. In this overview we give the description, unit of measure, the frequency it is measured and indicate whether it is controlled using a target.

2.3   Current required performance level
What are the current required performance levels on these KPIs?

To manage the complex turnaround process and provide Service Level Agreements (SLAs) with its suppliers, KLC has established fifteen turnaround schedules that define which activities have to take place at what time during a turnaround. These fifteen different types are based on certain circumstances in which the turnaround has to be performed. These include the type of aircraft, which gate the passengers have to pass through, whether it is a quick turnaround (QTA) or not, and whether it is the first flight of the day. Each of these circumstances has a certain influence on the exact characteristics of the turnaround; an Embraer needs more time than a Fokker, because it is a larger aircraft and some activities thus take longer (cleaning, boarding of passengers). In a QTA, there is less time available for certain activities. Embraer flights leaving at the D-gate might use three busses instead of two, and finally, the first flight of the day only needs departure-oriented activities. The fifteen different turnarounds with their respective parameters can be found in Table 4, they are internally documented as 'OCP-ZY20'. An example of one of these can be found in Appendix B.

Number	Aircraft type	QTA	Gate type	First of day?
1	F70	No	Semi contact	Yes
2	F70	No	Platform	Yes
3	F70	No	Semi contact	No
4	F70	No	Platform	No
5	F70	Yes	Semi contact	No
6	F70	Yes	Platform	No
7	E190	No	Semi contact	Yes
8	E190	No	Platform	Yes
9	E190	No	Platform, D-gate	Yes
10	E190	No	Semi contact	No
11	E190	No	Platform	No
12	E190	No	Platform, D-gate	No
13	E190	Yes	Semi contact	No
14	E190	Yes	Platform	No
15	E190	Yes	Platform, D-gate	No

Table 4 – Turnaround types at KLC

UNIVERSITEIT TWENTE.

Periodic & General				
General	**Metric**	**Frequency**	**Target**	**Description**
D0	Percentage	D, W, M, Q, Y	Yes	% of flights in that period that leave on or before scheduled time
D15	Percentage	D, W, M, Q, Y	Yes	% of flights in that period that leave within 15 minutes of scheduled time
A0	Percentage	D, W, M, Q, Y	Yes	% of flights in that period that arrive on or before scheduled time
A15	Percentage	D, W, M, Q, Y	Yes	% of flights in that period that arrive within 15 minutes of scheduled time
Completion factor	Percentage	D, W, M, Q, Y	Yes	% of scheduled flights which actually have been flown
Registration changes	Number	D, W, M, Q, Y	No	Number of changes to which aircraft is assigned to a flight within 60 minutes of departure
Non-performance departments	Hours	D, W, M, Q, Y	No	Sum of delay caused by each department of KLC, identified using delay codes
Non-performance within department	Hours	D, W, M, Q, Y	No	Sum of delay caused by each activity within a given department, identified using delay codes
Additional Department specific KPI's				
Ground Services				
Service providers ready at their deadline	Boolean	M	Yes	Whether the given service was finished at its designated moment
Additional specific measurements with different service providers, **Example measurement with KE (fuelling service provider):**				
# Delays caused by fuelling	Number	M, Y	Yes	Count of delay codes assigned to fuelling
# Refuels	Number	M, Y	No	Count of total refuels in period
% Refuelling's started before D-45	Percentage	M, Y	No	Percentage of refuelling's started before D-45
# Defuels	Number	M, Y	Yes	Count of total defuels (e.g. due to cancelled flight)
# Additional fuel	Number	M, Y	Yes	Count of additional refuelling's (e.g. due to changed weather conditions)
# Additional fuel after D-20	Number	M, Y	Yes	Count of additional fuel starts after D-20
% Refuels with >100KG overshoot	Percentage	M, Y	No	Percentage of refuels where the amount of loaded fuel was over 100 KG more than requested
Flight Operations				
Flight Plan / Fuel	Hours	M, Y	No	Sum of delay caused around flight- & fuel plan approval, code 57, 61 & 62 (Flight plan N/A, Flight plan change, addition fuel required)
Co Crew	Hours	M, Y	No	Sum of delay caused by Cockpit crew, code 63,64,65, 69 (late crew boarding, crew missing, crew special request, captain requested security check)

Table 5 - Current KLC Key Performance Indicators   *D = Daily, W = Weekly, M = Monthly, Q = Quarterly, Y = Yearly*

UNIVERSITEIT TWENTE.   KLM

2.3.1 Schiphol gate and aircraft position layout

The Gate type mentioned in Table 4 describes where the aircraft is located and where the scanning of the boarding passes of the passengers takes place. The aircraft can either be parked at the B-pier or on the platform, passengers can either be scanned at the B-pier or the D6-gate. A 'semi contact' departure means that the aircraft is parked at the B-pier and passengers are scanned at this pier as well. This allows passengers to, after their boarding passes are scanned, walk directly to the aircraft via two stairs without having to be transported by bus. A platform departure means that the aircraft is not parked directly at the gate and that passengers have to be transported to the aircraft first. For these departures, passengers can be scanned at either the B-pier and the D6-gate. After scanning, they are transported to the aircraft by bus. While most flights depart from the B-pier, this is not possible for all flights as some KLC destinations are not within the Schengen area, which means that movement of people and goods is not unrestricted. Examples of such destinations include the UK, Switserland, and Croatia. For these flights, an additional security clearance is needed, which Schiphol does not facilitate at the B-pier. Therefore, these flights have to depart from a special bus gate in the D-pier (D6-gate), where these facilities are offered. A section of the map of Schiphol where we marked the B-pier, D6-gate and platform can be found in Figure 4. Because of the larger distance, flights departing from the D6-gate with over 70 passengers are transported from the gate to the aircraft using three busses, all other flights (all B-pier departures and departures with less than 70 passengers from the D6 gate) use a standard of two busses. Other gates on Schiphol, such as those on the C-pier or the F-pier, are not used by KLC.
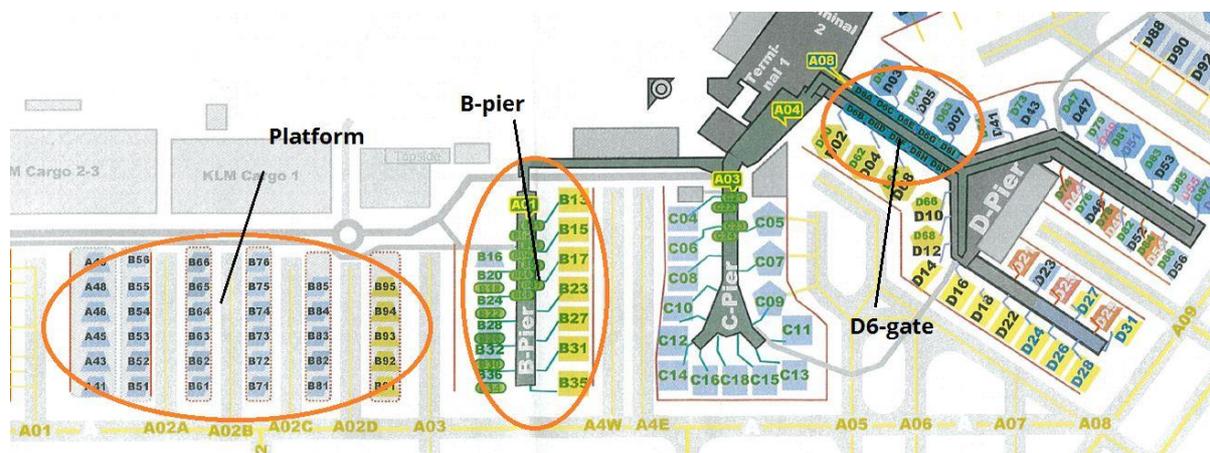
Figure 4 – Gate and aircraft positions at Schiphol

2.3.2 Scheduled ground time

A regular turnaround has a minimum ground time of 40 minutes for the Fokker F70 and 50 minutes for the Embraer E190. This distinction is made on the size of the aircraft. An Embraer can seat up to 100 passengers, compared to 80 for the Fokker. While a standard turnaround is often called a 40 minutes or 50 minutes turnaround, actual planned ground time can be far greater than that, as the mentioned 40 and 50 minutes are simply a minimum.

In Table 4, we mention the quick turnaround, or QTA. A QTA is a turnaround with a duration of 35 minutes, rather than the standard durations of 40 minutes, 50 minutes or longer. During such a turnaround, the activity time windows are minimized to a bare minimum, in order to get the aircraft back in the air again as fast as possible. These turnarounds can be used as a delay recovery tool (it can recover 15 minutes of delay when the original ground time is 50 minutes) or as a way to maximise aircraft utilisation during peak hours. To make sure they have the capacity and ability to service the aircraft within the given time limit, changing a regular turnaround into a QTA to recover from delay can only be done after consulting with the service providers. There is a

limit to the number of these unplanned and planned QTAs that KLC can request each period, so each request is weighted carefully.

Next to shortening the planned ground time, KLC has another option to handle delay, namely setting a new departure time. This new departure time, called an Estimated Block Departure (EBDE), can be set when the arriving flight is delayed for such an amount of time that the minimum ground time is violated. For example, an Embraer turnaround with a planned ground time of 60 minutes has an arriving flight with 20 minutes delay. KLC needs to set an EBDE at least 10 minutes past scheduled departure (STD), to allow all ground processes to finish their tasks properly. Important to note is that declaring an EBDE does not change the fact that the aircraft is delayed, any departure delay will still be calculated based on the original STD.

2.3.3 Service Level Agreements

KLC has set up a Service Level Agreement (SLA) with each of its service providers to ensure a reliable ground process. For this research we have analysed the SLA that KLC has agreed upon with their service provider on Schiphol, KLM Ground Services (KLM GS). In this SLA, or Standard Ground Handling Agreement (SGHA) as it is called, a wide range of agreements is gathered. These range from the percentage of flights that should be handled in time to the code of conduct for the driver who transports cockpit- and cabin crew to the aircraft before departure. This research focusses on the quantitative agreements regarding processes and punctuality and ignores all qualitative agreements.

Because the SGHA is an agreement between KLC and KLM GS that is valid for a period of 3 years and both parties have agreed that it is undesirable to have specific performance requirements set for such a long period, some of these requirements are determined in additional agreements with a much shorter time span. This enables KLC to adjust targets and requirements on a much shorter notice and makes them more dynamic. Nearly all quantitatively measurable agreements regarding performance of the various sub processes are captured in these additional agreements.

Regarding required quantitative performance levels relevant for this research, an interesting observation can be made. For individual activities, only the latest completion time is agreed upon, not the starting time or activity duration. The idea behind this is that as long the activity finishes in time while also adhering to the qualitative requirements, no problems arise. By not requiring a certain duration or starting time, the service provider is left with more flexibility in how it wants to fulfil its deadline, for example by scheduling additional workforce on a single aircraft to speed up the process.

As the entire turnaround process is managed by the service provider, additional agreements have been made on an overall turnaround level regarding punctuality of cargo-, passenger-, and "all doors closed" and D0 and D15 performance. These five performance requirements are of course intertwined but by measuring them all, more control over the sub processes is maintained. For example, when during a specific turnaround passenger boarding is delayed and thus the passenger door will not close in time, the cargo crew is still encouraged to close the cargo door in time to make sure the 'cargo door closed' measurement is still in time and the chance of additional delays occurring in this process is minimised.

2.4   Conclusion

KLC tries very hard to monitor the performance of the turnaround process at Schiphol. The entire process is well documented, frequent performance meetings take place and extensive service agreements with its service providers are in place. These service agreements have both a qualitative side, i.e. what should be performed, and a quantitative side, i.e. when it should be

UNIVERSITEIT TWENTE.

performed. Based on these agreements, KLC has created 15 turnaround process templates to make sure the activities are performed in time and do not interfere with each other. However, while these agreements mostly focus on activity completion time, KLC has little information that allows it to check the actual performance of the service providers. When looking at fuel for example, we see that KLC monitors the number of delay codes indicating a fuelling caused departure delay, but no data on actual numbers is used. In addition, most performance indicators are on a high level, indicating there 'is' departure punctuality non-performance, but possible causes are only identified using delay codes, which are subjective in their current form (see Section 1.2). We discuss the delay codes in more detail in Chapter 3, where we also gather the data that we need for our data analysis in Chapter 4.

UNIVERSITEIT TWENTE.

Chapter 3: Data gathering

This chapter answers the second research question: What data is further required to assess the actual performance of the turnaround process? To answer this question, we first look (Section 3.1) at the current delay identification process, how it is performed, and what its caveats are. In Section 3.2, we then turn our attention to currently known or suspected influences on the D0 performance by interviewing KLC employees. We try to identify which activities are currently underperforming or which circumstances negatively influence D0 according to KLC employees, and supplement this with findings from literature. In Section 3.3, we look at the data available for this research. As part of the AP2K project, a central data gathering system is being built which gathers info from all parties involved in the turnaround process. The data in this system is made fully available for this research. We discuss the range of data and what it represents. In Section 3.4, we filter our dataset to exclude unnecessary, incomplete, and inconsistent data, to make sure that our findings in Chapters 4 and 5 are relevant and reliable.

3.1 Current delay identification

How is delay identification currently performed?

As mentioned in Chapter 1, the current delay identification method is not waterproof. The codes are being assigned by hand by the Turnaround Coordinator (TLO) on duty; assignment is thus subject to the perception, opinion and knowledge of that specific TLO, resulting in not always properly assigned codes. For example, when the TLO notices that the boarding of passengers finished too late, he may assign code 15 (Boarding, Discrepancies, & Paging) as cause. However, the delayed boarding can also be caused by a preceding activity or other cause, e.g. code 67 (Cabin crew shortage), code 19 (Passenger with reduced mobility), code 17 (Catering Order Late/Incorrect) and many more. This requires a good overview of the entire process by the TLO and good communications between him and everybody involved in the turnaround process. However, given the often-hectic atmosphere during a delayed turnaround, this information may get lost, resulting in an incorrect delay code assignment. Another risk is that, given the large number of codes, the TLO simply does not know the right delay code to assign. Some codes are assigned more than others, resulting in the less used codes to be less known and maybe assigned even less. The final influence on the quality of the assigned delay codes is the number of delay codes that can be assigned. Currently only two codes can be assigned to a single flight, meaning that when there are more than two causes, this information gets lost. This means that while identifying causes of delay by using assigned codes is a good idea but in its current form its reliability depends on the skill, perception and honesty of the assigning TLO. This implies that although it is a good way to get a general impression of causes of delay, it is insufficient to gain true and decisive insights.

3.2 Known or suspected influences on D0 Performance

Which circumstances and activities are suspected or known to have a serious impact on D0 performance?

In this section, we try to identify which circumstances and activities are currently known or suspected to have a significant impact on D0 performance. These circumstances will be used as input variables in Chapter 4, where we test whether this information is correct. The primary source for these influences is internal knowledge at KLC supplemented with insights gained from literature. Special attention is given to the relation of departure punctuality to arrival punctuality of the preceding flight.

3.2.1 Propagated versus contributed delay

Before looking at circumstances or activities we mention a very important event, the actual time of arrival. This event enables all other activities and events to take place. When the arriving flight is

UNIVERSITEIT TWENTE.

delayed and there is not sufficient ground buffer time, all activities and the departing flight will also be late. However, this delayed flight could not start any sooner, so is not accountable for its own delay. Therefore, we introduce a two-type classification of delay: propagated delay and contributed delay.
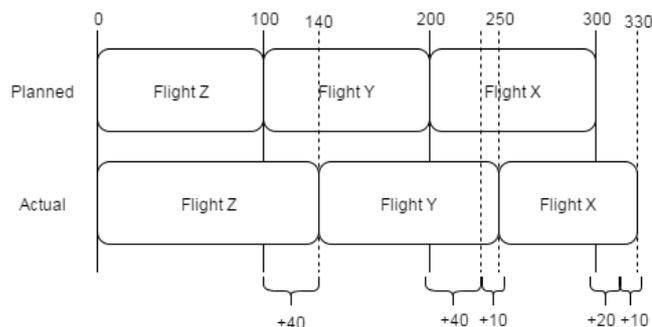

Figure 5 – Example of delay propagation

In this delay classification, propagated delay of a flight (say flight X) is delay caused by a preceding flight (flight Y) which could not be made up for by flight X (AhmadBeygi, Cohn, Guan, & Belobaba, 2008). A simple example: flight Y arrived 50 minutes late and, because there was no slack in the schedule, flight X also departs 50 minutes late. However, flight X manages to make up for some of the delay and arrives only 30 minutes past its original deadline. It would be unfair to hold flight X accountable for these 30 minutes delay as it even made up for some of the delay it received from flight Y. Therefore, we attribute the 30 minutes delay of flight X to flight Y. This delay propagation works on multiple levels; say we add a flight Z to the network, scheduled as predecessor of flight Y. flight Z started in time but finished 40 minutes late. In this case, the 50 minutes delay of flight Y were partly caused by flight Z being 40 minutes late, thus it would again be unfair to hold flight Y accountable for the full 50 minutes so we attribute 40 minutes to flight Z while keeping flight Y accountable for 10 minutes. Next to that, only 10 of the 30 minutes of delay of flight X are caused by flight Y. Therefore, the other 20 of these 30 minutes are in turn attributed to flight Z.

	Flight Z	Flight Y	Flight X
Planned start	0	100	200
Actual start	0	140	250
Planned finish	100	200	300
Actual finish	140	250	330
Delta throughput time	+40	+10	-20
Propagated delay	0	60	30
Contributed delay	100	20	0

Table 6 – Delay propagation calculation example

Contributed delay on the other hand is the sum of delays caused by a given flight, consisting of any delay in the flight itself that could not be propagated to preceding flights added with any propagated delay of succeeding flights (AhmadBeygi, Cohn, Guan, & Belobaba, 2008). Using the same example as above, we say the contributed delay of flight X is 0 minutes, as all of its delay could be propagated to flight Y. The contributed delay of flight Y is a bit more complex, but is equal to 20 minutes. It arrived 50 minutes late and received 30 minutes of propagated delay from flight X. Of these 50 minutes of own delay we propagate 40 minutes to flight Z for arriving late, thus leaving 10 minutes of his own delay caused by himself. Again, 20 of the 30 minutes of delay propagated by X are passed on to flight Z, leaving flight Y with having caused 20 minutes of total contributed delay. Finally, flight Z has a contributed delay of 100 minutes. It departed in time but

15

arrived 40 minutes late, thus making himself fully accountable for these 40 minutes. In addition, these 40 minutes of delay caused flight Y to be 40 minutes late and flight X to be 20 minutes late. We therefore say flight Z has caused 100 minutes of delay in the total chain.

This distinction between propagated and contributed delay enables us to better understand the origin of delay by accounting root causes for the full delay they cause. It also provides a good insight in how delay can 'echo' through a schedule when flights are scheduled with too little buffer time.

To attribute delay to activities within a turnaround, we use a slightly altered implementation. Because we are only interested in the delay of the final activity (departure), we do not propagate any delay of the other activities. This means that, when looking at the example in Table 6 as activities during a turnaround, only the 30-minute delay of flight X would be propagated, not the 50 minutes of Flight Y or the 40 minutes of Flight Z.

3.2.2 Circumstances

Circumstances have a large impact on departure punctuality (Wu & Caves, 2003). KLC knows this and, as mentioned in Table 4, has defined different turnaround types to match specific circumstances. Therefore, the variables KLC uses to define their turnaround types are used as input variables for expected influential circumstances. These include aircraft type, boarding gate pier and whether it is a quick turnaround and first of day or not. Of these variables, the one indicating whether it was a quick turnaround (QTA) will be slightly altered to accommodate a more detailed analysis. The turnaround types define a duration of either 35 minutes for QTAs or 40 and 50 minutes or standard turnarounds. However, these 40 and 50 minutes are a minimum duration, not accounting for any buffer time. In reality, a standard turnaround can be anywhere between 50 minutes and multiple hours. Therefore, we will use planned ground time as an input variable, rather than just the QTA indicator.

KLC operates a hub-and-spoke network, which means that it operates all flights from and to a single airport (Schiphol, the hub). Therefore, roughly 70% of all passengers have transfers between flights. Not only between KLC flights, but mostly with intercontinental flights of the mother company KLM and other partner airlines. In order to try and minimize total travel times for its passengers, transfer times are kept to a minimum, meaning most flights are served in two timespans, namely in the morning and early evening. This means that during these 'rush hours' the traffic intensity at the airport is expected to be high, passengers and luggage from other flights have to be transferred and available ground service crews are limited. Therefore, KLC expects the time of day to be an important influence in D0 performance. This is confirmed by (Hsiao & Hansen, 2006) where a strong influence of time of day was found on queuing length in flight plans.

Another circumstance which, according to both KLC and literature (Wu & Caves, 2003b), has a large influence on D0 performance is traffic intensity at the airport. When many flights are arriving and departing, Air Traffic Control (ATC) might deny the aircraft to be pushed back for departure due to congestion on run- and taxiways. As a turnaround is only completed when the pushback is complete, such a situation will lengthen turnaround duration, and thus deteriorate D0 performance. However, since we do not have any data regarding the total number of flights arriving at- and departing from Schiphol at any given moment in time, we will assume traffic intensity to be incorporated in the previously mentioned Time of Day.

Similar to Time of Day there are two additional variables that are used as input variable, namely Day of Week and Time of Year. The number of flights fluctuates between the different days in a week and between seasons. Therefore, longer-term congestion effects are measured using these two variables.

UNIVERSITEIT TWENTE.

An obvious circumstance to include is weather. Although en-route weather mainly influences flight- and arrival times, terminal weather (the weather at the airport) has some influence on the turnaround process itself. However, previous research (Hsiao & Hansen, 2006) suggests that this influence only becomes noticeable when there is high traffic intensity and even then only slightly. Still, we use data gathered from the KNMI (Dutch Weather Institute) to see whether there is an impact present. Given the focus of this research on departure- and not arrival punctuality (A0 performance), we ignore en-route weather.

An important circumstance is the presence of an unaccompanied minor (UM) or a passenger with reduced mobility (PRM). Although currently not tested with actual data, KLC suspects this to have an impact on departure punctuality. These passengers have to be transported separately to and from the aircraft and also (de)board separately. Especially the PRMs might have a strong influence as they often need a long time getting up and down the stairs used to board KLC aircraft at AMS. The variable used for this test is the presence of an Axxicom bus during the turnaround as the company Axxicom is responsible for transporting the UM or PRM to and from the aircraft.

In Section 3.2.1, we discussed the importance of delay propagation and how delay of one flight can cause the next flight to also delay. Long delays on preceding flights can, together with other factors such as crew sickness or technical issues, be a reason to assign a flight to a different aircraft. During the day, KLC has multiple aircraft in reserve, precisely for this purpose. In addition, in such situations flights can also be swapped between different aircraft when this reduces total amount of delay. Changing a flight to be performed by a different aircraft is called a registration change and can be very useful in reducing the impact of delay on the total schedule to a minimum. However, a late registration change can require a lot of work. For example, a new crew might has to be assembled, when the gate changes this has to be communicated with both customers and the service providers, additional busses may have to be requested at the airport. Therefore, a registration change within an hour of scheduled departure is very inconvenient for all involved parties and while it may reduce total delay, the flight will almost certainly depart with a delay.

3.2.3 Activities
Next to the general circumstances discussed in Section 3.2.2, additional delay can arise at activities during the actual turnaround process. This can be due to poor scheduling, poor performance or other reasons. However, as detailed information regarding ground service performance was unavailable for this research, little preliminary information can be given here. We expect that activities that lie on the general critical path contribute most to departure delay. Literature suggests that these activities are de-boarding of arriving passengers, cleaning and catering, boarding of passengers and pushback. By applying delay propagation on the data and the OCP-ZY20 procedure, we test this hypothesis.

One 'activity' expected to have a significant influence on departure punctuality is waiting for departure clearance from local air traffic control (ATC). This is confirmed by both literature (Tu & Ball, 2008), and KLC expectations. When the aircraft is ready to depart, the captain requests a departing slot from ATC. The time between this request and the actual start of the pushback can, given the pushback truck was ready, be attributed to local ATC. As discussed in Section 3.1, code 89 (Local ATC) is by far the code that is assigned most when a flight is delayed. Using the timestamp of the 'ready to depart' notification, we will try to verify whether this assignment is justified.

3.3 Data availability of historical performance
What data is available regarding the historical performance of activities during a turnaround?

The amount of data generated in the airline industry is staggering. A single flight creates thousands of measuring points. All these data points are stored in different systems of various parties involved and range from the scanning time of the boarding pass of each individual passenger or the moment the aircraft's cargo door is opened to the time that a bus transporting passengers to the aircraft is on position or the cleaning crew is ready. In fall 2014, KLC started a project titled 'Analysis Process Performance KLC' or AP2K. The aim of this project was to provide insight in D0 performance by gathering and combining relevant data from the various parties involved in a turnaround.

Over the course of this research, the AP2K project developed a proof-of-concept where they went from a single data dump of data towards a daily exchange of data, enabling short-term assessment of a turnaround. For this research, a selection of the already present data was used, together with additional data that, based on the preliminary research of Section 3.2, might prove to have an influence on departure punctuality.

3.3.1 Available data types

As mentioned, thousands of measurements are recorded each flight, ranging from trivial events to milestone events such as the "all doors closed" (ADC) moment logged in the aircraft's internal logging system. Dependent on the source and type of data, these timestamps are either system generated and thus hard, or manually given and thus soft. Hard timestamps are undisputable as the generated timestamp exactly corresponds with the actual time of the event. Soft timestamps on the other hand can be open for discussion. For example, the cleaning company issues handhelds to their workers in which they have to indicate when they start or end the cleaning of a specific aircraft. Based on these check-in and check–out times, among other uses, regular performance assessment is performed to check each individual worker's performance. However, the exact moment 'done with cleaning' is ambiguous as it can be interpreted in multiple ways. For example, it can be the moment the workers are done but still in the back of the aircraft, it can be the moment they leave the aircraft, or it can be the moment they are back in their car and ready to leave for a new job. It is in the interest of the worker to end his task at the first example, as this will result in the shortest total cleaning duration and thus a better performance. However, when determining departure delay causes it might skew results as the data suggests cleaning is done and passengers can board the aircraft while in reality the cleaning is still aboard and passengers still have to wait. KLM Ground Services continuously tries to improve the reliability of these soft measuring points. For this research, we therefore assume these measuring points to be an approximation rather than a definite timestamp.

Important to note is that some data points used in this research are deducted from other data points. A good example of this is propagated and contributed delay. Using the method presented in Section 3.2.1, these values were calculated for each turnaround and event during a turnaround. This calculation was already executed by KLM as part of the AP2K project and was used as input for this research. Other calculated variables are straightforward data types such as activity duration and time between activity completion and scheduled time of departure. Table 7 presents all variables at our disposal, together with their unit of measure.

3.3.2 Data sources

As mentioned, the AP2K project uses different data sources to provide an integral overview of the turnaround process. To give some idea on where our data originates, we list the various sources we use, either via the AP2K project, or directly. The most important data sources are:

UNIVERSITEIT TWENTE.

The HubDB is a database with vast amounts of data generated by KLM Ground Services, supplemented with various sources relevant for the ground process. Most of the sources mentioned hereafter are captured within the HubDB. However, because we believe information regarding the origin of data is useful we introduce the most important indirect sources individually.

AAS

AAS (Amsterdam Airport Schiphol), is responsible for the busses transporting passengers from the gate to the aircraft when it is a so-called platform departure. The drivers of these busses generate various data points, such as the time they are on position and waiting for passengers, the time the offloading of the passengers starts at the aircraft and the time offloading is completed. In this research, this data is used as an indicator for the performance of the passenger boarding process, as there is currently no other data available besides the scanning of the boarding passes at the gate. Unfortunately, the timestamps are generated manually by the driver of the bus and there might be a discrepancy between the actual event and the logged timestamp.

ACARS

ACARS (Aircraft Communications and Reporting System), is a communication system linking aircraft to ground control. One of its uses, and the most relevant one for this research, is the logging of all system events generated by the aircraft before, during, and after a flight. Data gathered includes 'first cargo door open', 'last cargo door closed', 'passenger door open', and 'passenger door closed'. Because this data is generated by the aircraft itself, it is highly reliable. Data from ACARS is gathered from the HubDB.

Axxicom

Axxicom is the company responsible for transporting unaccompanied minors (UM's) and passengers with reduced mobility (PRM) from and to the aircraft. Information gathered from this source includes collection and drop-off time of the passengers. The timestamps we received from Axxicom are unfortunately also generated by hand.

CHIP

CHIP, (Dutch: Communicatie & Hub Indelings-Programma) or Communication and Hub Scheduling programme, is the system KLM Ground Services uses for online operations scheduling of most tasks; TLO, cleaning, fuelling, cargo, toilet, water, and pushback. It has a constant feed of data of which tasks have to be performed at what gate and creates a load balanced schedule based on both priority and urgency of each task. Of each activity the planned start and end moment are registered, together with the actual start and end moments. Important to note is that the planned data is system generated while the moments at which the task was actually performed is generated manually by the operator. All CHIP data is gathered from the HubDB.

FIRDA

FIRDA, (Flight Information Royal Dutch Airlines), is the system in which flight data is recorded for every flight executed by KLM. Examples of data points include scheduled time of arrival (STA), actual time of arrival (ATA), and Estimate Block Departure (EBDE, adjusted scheduled departure time to accommodate circumstances such as a late incoming aircraft). Data from FIRDA is system generated and gathered from the HubDB.

LVNL

LVNL, the Dutch air traffic control, is the source of one specific measurement point: the timestamp of the 'ready to depart' notification from the cockpit to the tower. This notification should be given when everybody is seated, all checks are complete and the aircraft is ready for departure. This

measurement can be useful in differentiating between delays originating from local ATC or from the aircraft not being ready, for example because not all passengers are seated yet. This measurement is system generated and thus reliable.

KNMI

KNMI, the Dutch weather institute, is the source of our weather data at Schiphol. The KNMI provides a historical per hour overview of the raw data of each of its measuring instruments at each of their stations. To get the weather conditions at Schiphol during time of departure, we process this information from the Schiphol weather station into a single classification with the following options: thunder, snow, fog, heavy rain, rain, and other. Together with the average wind speed and wind direction we match this hourly weather information with the scheduled departure time of the flights.

3.3.3 Full data list

We have created an overview of data that is used as input for our analysis in Table 7. The data is grouped, labelled whether it was system generated, manually generated or calculated, described in unit of measure and a description is given. The additionally calculated values are fairly self-explanatory, but to prevent confusion, we discuss some of them briefly.

Planned ground time is determined by the time difference between scheduled time of arrival (STA) and scheduled time of departure (STD), while the real ground time is determined by the time difference between actual time of arrival (ATA), and actual time of departure (ATD). The 'registration change in time' Boolean indicates whether the latest registration change was done early enough to meet KLC's own requirements. As a registration change requires substantial changes to the workflow, it should be communicated at least one hour before scheduled departure. The activity- and event lateness and 'activity late start' minutes are determined by calculating the time at which the activity should have taken place according to KLC's own OCP-ZY20 procedure and determining the difference with the actual times recorded. This can be used to check the conformance to the OCP-ZY20 procedure template. These times are also calculated for the ready to depart and actual departure moments, to see when the aircraft was actually ready and how long it still had to wait for ATC clearance.

Name	Type of measurement	Unit of measure
General turnaround information *For each turnaround we have the following data*		
Date	Descriptive	Date
Aircraft Registration	Descriptive	Descriptive
Aircraft Type	Descriptive	Descriptive
Station	Descriptive	Descriptive
Inbound Flight Number & Station	Descriptive	Descriptive
Inbound Gate Number	Descriptive	Descriptive
Outbound Flight Number & Station	Descriptive	Descriptive
Outbound Gate Number	Descriptive	Descriptive
Outbound Load factor	Calculated	Percentage
Weather	System generated	Descriptive
Wind speed & direction	System generated	Descriptive
Table continues on next page		

UNIVERSITEIT TWENTE.

General ground time information		
For each turnaround we have the following additional data		
Scheduled time of arrival (STA)	System generated	Date + Time
Scheduled time of departure (STD)	System generated	Date + Time
Estimate block time (EBDE)	System generated	Date + Time
Estimate block time NOC (EBDL)	System generated	Date + Time
Number of registration changes	System generated	Integer
Latest Registration change	System generated	Date + Time
Registration change in time	Calculated	Boolean
Minimum ground time (MinSTD)	Calculated	Date + Time
Planned ground time	Calculated	Minutes
Real ground time	Calculated	Minutes
Delay information		
For each turnaround we have the following delay data		
Delay code 1	Manually generated	Descriptive
Delayminutes1	Manually generated	Minutes
Delay code 2	Manually generated	Descriptive
Delayminutes2	Manually generated	Minutes
A0	Calculated	Minutes
D0	Calculated	Minutes
ContributedD0	Calculated	Minutes
PropagatedD0	Calculated	Minutes
Events*		
For each of the events that happened during a turnaround we have the following data:		
Event Time	System generated	Date + Time
Event Lateness	Calculated	Date + Time
Event Dminus	Calculated	Date + Time
Event Aplus	Calculated	Date + Time
Regular activities**		
For each of the activities performed during a turnaround we have the following data:		
Activity start	Manually generated	Date + Time
Activity end	Manually generated	Date + Time
Activity duration	Calculated	Minutes
Activity late start	Calculated	Minutes
Activity lateness	Calculated	Minutes
Activity D-minus	Calculated	Minutes
Activity A-plus	Calculated	Minutes

Table 7 – Available variables per turnaround

Actual time of arrival (ATA), Actual time of departure (ATD), Passenger door open (PDO), passenger door closed (PDC), cargo door open (CDO), cargo door closed (CDC), all doors open (ADO), all doors closed (ADC) and Ready to Depart (RTD).
**Axxicom, Bus, Catering, Crew Transport, Cleaning, Fuel, ExtraFuel, PaxScan Pushback, TLO, Toilet, Water*

UNIVERSITEIT TWENTE.

3.4 Distilled dataset

Which data can be used for the performance assessment?

The available data for this research was far from complete or unambiguous. Because not all service providers were able or willing to share their information, it proved to be hard to gather a full and complete set of data points. Effort was made to gather data ranging from 01-01-2014 until 01-06-2015. However, we did not succeed for all data providers. In Figure 6, an overview of all data sources and their availability are presented. As we see, most data is available for the entire timespan. However, Axxicom, AAS, and LVNL data is partially missing from our dataset. Unfortunately, we were unable to gather additional data from these sources.

3.4.1 Exclusion rules

When using empirical data for an analysis, one must make sure that the observed data is reliable and validated (Flynn, Sakakibara, & Schroeder, 1990). In our analysis, this means that we need to exclude any obviously incorrect or irrelevant data to improve reliability of the results. For this purpose, we set up a set of exclusion rules that we applied to the data. These rules were verified and justified by KLC to make sure we did not exclude relevant data. These rules are:

Each outbound flight should be a regular flight with passengers

Not all of our data concerns regular turnarounds. Sometimes an aircraft cannot fly due to technical failures or damages and a replacement aircraft has to be used to perform the flight. If this happens at an outstation, often no replacement aircraft is present and it has to be flown in from AMS. This 'aircraft placement flight' has no passengers on board and is logged with a special flight number to differentiate it from regular flights. Because no full servicing is needed and departure delay is not of influence to total KLC numbers, these flights should not be taken into consideration.

Each outbound flight should be to a regular destination

Next to aircraft placement flights, we also exclude non-regular destinations. Some flights have a destination that is not in the regular schedule. This includes flights that are normally flown by the mother company KLM, but due to circumstances are flown by KLC. These flights are incidental and because of the irregular nature of these flights, KLC has indicated that these turnarounds should be ignored.

Figure 6 – Data availability per source

Each outbound flight should be no more than three hours delayed

A delay of over 3 hours is unlikely to be caused by regular delay causes. Often it is a disruptive cause that cannot be anticipated or circumvented, e.g. the closing of the airport due to a terrorist threat or a technical malfunction. Because we have no information regarding the occurrence of these exceptional causes and their predictability is low, we eliminate these turnarounds from our data.

Each outbound flight should not depart more than sixty minutes early

An aircraft can only depart when all passengers are present and boarded, which is unlikely to be the case more than an hour before scheduled departure. Therefore, a departure of more than 60 minutes early indicates special circumstances or a data error. Both are unwanted in our dataset.

The total planned turnaround time should be non-negative and non-zero

The planned turnaround time is the time difference between the scheduled time of arrival (STA) and scheduled time of departure (STD). When this time is negative, i.e. the aircraft is scheduled to depart before it is scheduled to arrive, this indicates a data error and we eliminate it from our dataset. Similar, when either STA or STD is missing, planned turnaround time cannot be determined and the turnaround is eliminated from the dataset.

All measurements of all major activities should be present

To get a thorough and reliable understanding of delay causes, it is necessary to have consistent measurements. When parts of the data of a single turnaround are missing, no reliable conclusion can be drawn as delay might be caused by a missing variable. Therefore, we have limited our dataset to include only turnarounds for which the measurements of all major activities are present. These activities include baggage, fuel, toilet, catering, cleaning, crew transport, and boarding pass scans. Some activities are left out of this list because they do not necessarily occur every turnaround. Busses are only needed when departing from the platform, Axxicom is only needed when there is a PRM or UM present, pushback is not always needed as some parking spots can be exited in a forward movement and water is sometimes multi stretched, meaning multiple flights are performed on a single load.

All activity measurements should be within logical boundaries.

To exclude obvious data errors, we eliminated all turnarounds for which any activity had a timestamp outside logically possible boundaries. For example, when the operator of the toilet cleaning truck 'completes' his task after the aircraft departed, it is an obvious error. We decided not to place these boundaries exactly at actual arrival and departure, because sometimes servicers start their activity when they start to drive towards the location of the aircraft, thus 'starting' before the aircraft actually has arrived. For example, the Baggage handling operators start their activity when they start loading baggage onto their trucks at the baggage collection point. At departure, we have the same situation. A clear example is the pushback operator. He finishes his task when the aircraft is pushed back, thus after the actual departure (which is equal to the moment he starts pushing). We set our boundaries differently for each of the major activities to make sure they are not too loose for some activities, or too restrictive for others.

All major activities must be longer than two minutes in duration but no more than sixty.

First, we eliminate any activities that have a negative duration, and thus contain a data error. Next, by cutting at two minutes (instead of the more obvious one minute), we eliminate skewness in the data because operators start / end their task in the same minute. This might be because he finds out he forgot to start his activity when he wants to stop it, because he did not want to log his duration, or many other reasons. Exception to this rule is the water service. As servicing the aircraft can take sometimes less than a minute when only a small top-off is needed, we have

required this activity's duration to be at least one minute. Note that activity durations are always rounded upwards to integer values.

We only set a limit on the length of the duration if the turnaround is not an overnight stay. This is because some activities during an overnight turnaround require multiple operations from a single service provider, one in the evening and one in the morning. In our data, this is represented as a 'start activity' timestamp in the evening and an 'end activity' timestamp in the morning, resulting in an extreme duration of that specific activity. Unfortunately, we cannot reproduce the original data of both separate operations. However, if we choose to eliminate all turnarounds with this kind of data, we eliminate all 'first of day' turnarounds. Given their importance to the entire operation, we chose to keep these turnarounds into our dataset.

Removing variables

In addition, we exclude certain variables from our data set that might influence our analysis negatively. For example, the D0 performance (indicator of the amount of delay) and any assigned delay codes would provide our data analysis in Chapter 4 with a perfect predictor of delay and should therefore not be included. Also, the time of completion of the pushback task is unwanted, as it is always after departure, thus a perfect predictor. All variables we remove are listed in Table 8. Note that while we remove the variables concerning the completion of the pushback task, we keep the variables concerning the starting of the pushback task. While it can only finish after departure, it can still start its task by connecting to the aircraft when other tasks are still being performed.

Variable	Unwanted influence
Delay 1 & Delay 2	A delay code is only present when the flight is delayed
Delayminutes1 & 2	Delay minutes are only present when the flight is delayed
D0	Being the time between STD and ATD, it is a perfect 'predictor' of delay
All specific time events	We converted everything to relative times, dates are incorporated into month
Pushback-Dminus	The pushback task is only completed after ATD, so time between it and STD is again a perfect predictor of delay. PushbackStart-Dminus is still usable
Pushback duration	Closely related to Pushback End and Pushback-Dminus

Table 8 - Excluded variables

3.4.2 Resulting dataset

Over the entire timespan from the first of January 2014 until the first of June 2015, data from 138,694 turnarounds was gathered. Of these, 70,099 were at Schiphol. When applying our exclusion rules we are left with 47,120 turnarounds, roughly 67% of the available data. Table 10 shows an overview of how many turnarounds each rules eliminates. We decided not to split the 22,781 turnarounds eliminated by activity related measurements into the different rules, because many turnarounds are dropped for multiple reasons. Table 9 lists quality of our data with respect to the activity related measurements. Note that the 'total faulty' column indicates the number of turnarounds where at least one of the rules is violated, if a specific turnaround has multiple data errors for that activity, it will still only count as one faulty turnaround. For example, activities with out of bound measurements often also have an extreme duration. Special notice should be given to the catering task; we have a significant amount of turnarounds (roughly 10% of our data set) without catering data, although each KLC flight is serviced by the catering service. The reason why this data is missing is unknown.

UNIVERSITEIT TWENTE.

Activity	Turnarounds	Total Faulty*	Missing	Out of Bounds	Short Duration	Extreme Duration
Bax **	69,901	1,954	642	1,081	105	214
Cleaning **	69,901	1,792	462	346	170	937
Crew Transport	69,901	17,260	1,408	960	1,931	13,632
Fuel **	69,901	2,234	407	956	114	874
Extra Fuel	69,901	67,626	67,574	41	15	-
Toilet **	69,901	2,240	597	521	852	429
Water	69,901	18,099	17,444	143	-	536
Catering **	69,901	11,468	7,187	15	1,459	2,811
PaxScan **	69,901	3,788	304	2,658	811	2,336
TLO	69,901	45,935	197	12,464	286	42,936
Pushback	69,901	26,827	24,069	154	2,580	80
Arrival Bus	69,901	53,202	51,635	276	1,308	2
Departure Bus	69,901	52,214	52,149	56	6	3
Arrival Axxicom	69,901	62,419	60,879	663	1,197	11
Departure Axxicom	69,901	65,214	64,141	21	6	1,060
Event						
CDO	69,901	31,167	31,127	40	N/A	N/A
CDC	69,901	28,101	28,089	12	N/A	N/A
PDO	69,901	28,294	28,251	43	N/A	N/A
PDC	69,901	28,015	28,005	10	N/A	N/A

Table 9 – Activity data quality

Number of turnarounds with at least one faulty or missing measurement
**Activities that need to be present and not faulty*

3.5 Conclusion

In this chapter we have identified the current way KLC performed delay identification, which data we have available for our research and how we have filtered this data to gain a dataset to perform a reliable analysis on. This resulted in roughly one third of our data to be excluded from analysis. While this is unfortunate for our research, it shows that the data quality of the manually generated data is poor, and improvements in this data could increase the relevance of this research greatly. The result of our data gathering and data cleaning process is a set of 47,120 turnarounds for which we have a valid entry on all variables at our disposal (Table 7). In Chapter 4, we use data mining and statistical techniques to identify which of these variables have a significant impact on departure punctuality.

UNIVERSITEIT TWENTE.

KLM

Rule	Nr of dropped turnarounds	Turnarounds left in dataset
Start	-	70,099
Outbound flight should be a regular flight with passengers	103	69,996
Outbound flight should be to a regular destination	15	69,981
Total planned turnaround time should be non-negative	33	69,948
Outbound flight should be no more than 3 hours delayed	46	69,902
Outbound flight should not depart more than 60 minutes early	1	69,901
All measurements of all major activities should be present, within bounds and a have logical duration	22,781	47,120
Result		47,120

Table 10 – Data cleaning process and size of the resulting dataset

UNIVERSITEIT TWENTE.

Chapter 4: Current Performance Analysis

In this chapter, we discuss the results from the analysis into departure delay. We first discuss the general performance regarding duration, start and finish times of the activities during a turnaround (Section 4.1), after which we introduce decision tree learning and perform this on our data to find relations between circumstances or activity characteristics and departure delay (Section 4.2). In Section 4.3, we compare our delay classification to the current delay coding system to see which of these two gives a more reliable indication of the cause of delay.

4.1 General performance
What is the performance of the various activities?

In this section, we discuss the performance of the various activities during a turnaround by looking at duration (Section 4.1.1), completion time (Section 4.1.2), and starting time (Section 4.1.3). Note that for the completion time and starting time, we use the variables 'Dminus' and 'Aplus' that we determined from our data. The Dminus variable is the time between the completion of the task and a modified departure time. This departure time is the latest of any of three variables: Scheduled time of departure (STD), Estimated Block time (EBDE) and the minimum ground time. The latter was added to ensure service providers receive enough time to perform their activities, even when KLC did not set an EBDE to account for arrival delay. The Aplus variable is the time between the activity start moment and actual arrival, a performance measurement useful for
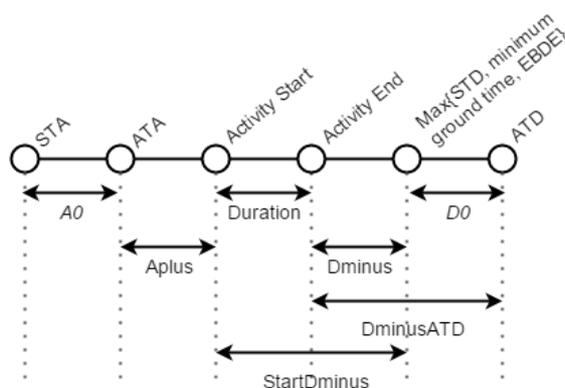
Figure 7 – Calculated variables based on available data

arrival-oriented activities. Figure 7 visualises these two variables, along with a few others. Note that arrival oriented and departure oriented targets and performance is expressed as 'A+#' or 'D-#', where '#' can be any number, e.g. A+3 or D-3. This indicates that the target for that activity is set as '3 minutes after arrival' or '3 minutes before departure'. Consequently, when reading the tables in Section 4.1, note that an activity with a target of A+3 and a performance of 4 (or A+4) is under performing, as it is after target. At the same time, an activity with a target of D-3 and a performance of 4 (or D-4) is over performing, as it is before target. Note that due to confidentiality the numbers in this section are altered and cannot be used as a reference for actual performance.

4.1.1 Activity duration

As mentioned earlier in this report, there is no fixed activity duration agreed upon by KLC with its different service providers. The focus in the SLA is mainly on qualitative factors and the time at which the activity is completed. Of course, this often implies a certain duration limit, as the various services can only start at a given moment during the turnaround and the performed activities need a specific amount of time. However, most activities have slack, as the available time window is larger than the time needed for the activity, or the activity duration is flexible. For example,

UNIVERSITEIT TWENTE.

cleaning can be performed by three instead of two workers, shortening the activity duration slightly. Thus, KLC has no fixed activity duration norms and leaves this to its contractors.

For comparing the activity duration performance, we therefore use the planned activity times in CHIP when available. As mentioned earlier, CHIP is used by KLM Ground Services (KLM GS) for their own operational planning and worker assignment. In CHIP, workers get a specific target time in which they have to perform their task. Of course, KLM GS has tried to match this assigned target to the actual workload, to obtain schedules in which workload is properly balanced and enough workers are available to perform all tasks at any given moment during the day. Therefore, the target activity times in CHIP are a good benchmark to test actual activity duration. An exception to this rule is the crew transport, the department responsible for this activity uses the CHIP planned times as a placeholder and actual times are internally allowed to vary. For all activities that are not planned in CHIP we used the OCP-ZY20 target start and end times to determine target duration times.

Activity	Target duration (minutes)	Actual average duration (<60 minutes)	Standard deviation of Actual duration
Arrival Bus*	6	7.44	2.19
Departure Bus*			
D6-gate	20	19.35	5.18
Platform	18	18.18	5.19
Arrival Axxicom	-	12.88	7.63
Departure Axxicom	-	31.15	12.76
Baggage handling	35	20.09	9.46
Catering	8	8.89	5.21
Cleaning	9	10.26	3.00
Crew Transport	12	21.61	14.90
Fuelling	14 - 21	15.85	5.71
Passenger Scans*			
Platform	15	13.74	4.89
D6-gate	19	15.52	4.55
Semi-contact	11	9.12	5.35
Pushback	4	3.96	12.30
Toilet	6	4.08	3.33
Water	5	2.09	2.95

OCP–ZY20 target

Table 11 - Target vs actual activity durations

All target and actual activity times can be found in Table 11. When multiple targets are listed, it means that different turnaround types require a different duration. Fuelling has a target range between 14 and 21 minutes, this is due to the scheduled duration to be dependent on the amount of requested fuel. The amount of requested fuel is not dependent on turnaround type but on many factors, such as weather and destination. Because a distinction between these is not relevant for this research, and to improve readability, we decided to group them together into a single performance measure. Next, Axxicom does not have specific target durations. For example, when arriving, Axxicom can drop off the passenger anywhere, depending on the needs for that specific passenger. The duration can vary greatly, depending on whether the passenger is transferring on AMS, in which case the location of the other aircraft can make a big difference. As long as the pick-up was done in time, its moment of completion is of no importance to KLC.

UNIVERSITEIT TWENTE.

Looking at Table 11, most activities seem to perform very well with an average duration that is similar to or smaller than its target. One activity stands out negatively: crew transport has an average duration of 21.61 minutes versus a target of 12 minutes. The explanation for this lies in the fact that the department responsible for crew transport does not use its planned CHIP times for anything other than to assign a job to a driver and thus not as a target. Because the actual duration can vary greatly depending on where the crew has to be picked up and delivered, the standard deviation of 14.90 minutes is also high. Note that we excluded durations of over 60 minutes in this overview. This is to prevent first of day turnarounds with multiple activities to skew the results. We discuss this in more detail in Section 4.1.3.

4.1.2 Activity completion time

In order to measure the performance of the various activities with regard to their completion time, we have to perform two steps. First, we have to make sure we measure the right performance and in order to do this, we determine the moment we want to compare the activity completion time with. Comparing it to the Actual time of departure (ATD) is not a good idea because when a flight is delayed, activities that caused the delay will seem to perform normal, while activities in that same turnaround that finished in time will seem to perform better than expected. Therefore, we set the baseline to be either of three moments in time: scheduled time of departure (STD), minimum ground time departure (minSTD) or Estimate Block Departure (EBDE). The STD is the originally scheduled time the aircraft should depart and is always present in our data. The minSTD is the time at which the minimum turnaround time (35 minutes, 40 minutes or 50 minutes) is passed after arrival. This is useful when the incoming flight is delayed, as operators still need their minimum time to complete their task. The EBDE is an adjusted scheduled departure time when an aircraft cannot depart sooner due to external reasons. This might for example be due to a severely delayed incoming flight, technical difficulties or restrictions at the airport. Of these three times, we take the one that represents the latest moment in time and determine the difference in minutes between that time and the activity completion time. This number of minutes before departure is stored in the variable D-minus and compared to the target completion time in Table 12.
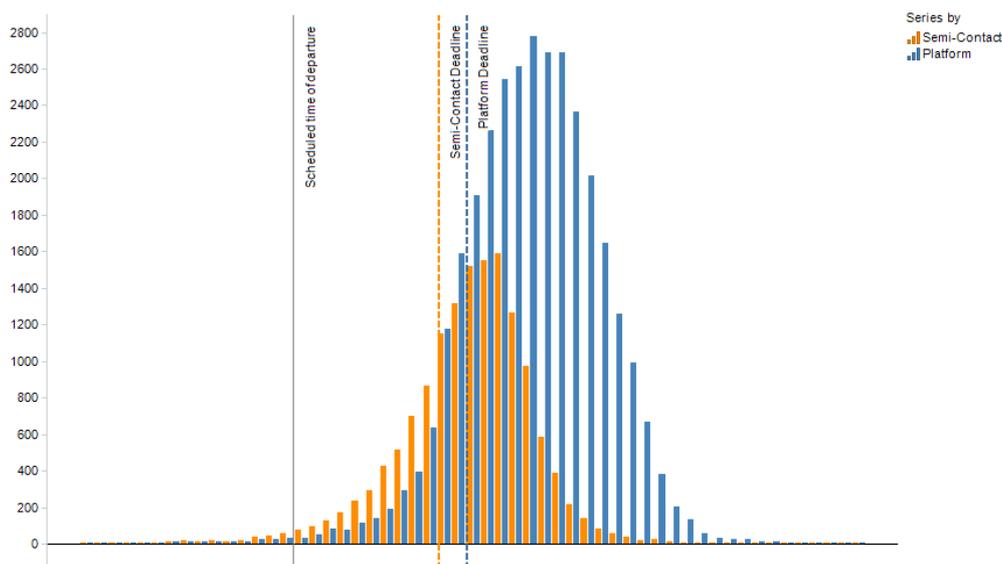
Figure 8 – Passenger scans completion time, platform and semi-contact

For each turnaround type with a different target, a separate listing is created in Table 12. When multiple turnaround types have the same target, but their performance differs significantly, it is listed separately. This is most often true for quick turnarounds (35 minutes) versus standard turnarounds. For example, cleaning has the same target completion time during a QTA and a

standard turnaround at the Semi-Contact gate (both 21 minutes before departure). However, the average completion time differs greatly (29 minutes versus 40 minutes). This is because the aircraft can be serviced anytime between arrival and departure, and longer ground time means a wider spread, resulting in a different average completion time.

Activity	Target	Actual average D-minus (end)	Standard deviation D-minus (end)
Departure Bus	D-7	9.31	4.61
Departure Axxicom	D-14	17.39	15.02
Baggage handling	D-1		
Catering			
35 min Platform	D-17	21.93	8.66
35 min Semi-Contact	D-21	23.23	8.34
Other F70	D-19	48.14	24.01
Other E90	D-25	46.36	22.19
Cleaning			
35 min Platform	D-18	23.18	6.54
Other Platform	D-18	48.06	30.47
35 min Semi-Contact	D-21	29.15	6.44
Other Semi-Contact	D-21	40.38	23.67
Crew Transport	D-30	39.78	
Fuelling			
35 min	D-6	7.56	9.80
Other	D-6	21.73	28.17
Passenger Scans			
Semi-Contact	D-8	9.10	4.86
Platform	D-9	11.42	5.41
Toilet			
35 min	D-23	31.15	10.84
Other	A+25	-	-
Water			
35 min	D-20	29.74	18.03
Other	D-20	44.33	116.18

Table 12 – Activity completion times compared to OCP-ZY20 target

Note that some activities have arrival-oriented targets, such as the busses picking up the arriving passengers. These activities are irrelevant for departure delay and are not included in this table. Special attention goes to the toilet servicing. In case of a standard turnaround, the deadline is set to be 25 minutes after arrival. This is to prevent the toilet from not being serviced for a longer timespan when the aircraft is not needed again right away, which is unwanted from a hygiene perspective. During QTAs, the toilet has a departure-oriented target and therefore is included in the table.

Looking at Table 12, the average completion time of each of the activities is quite good, with few of them having an average close to the deadline but most are completed a long time in advance. However, when accounting for standard deviation, we see a few activities that will regularly exceed their deadline. For example, passenger scans at a Semi-Contact gate are completed at D-9 on average, but with a standard deviation of nearly 5 minutes, its deadline of D-8 will be exceeded regularly. This is illustrated in Figure 8 (page 29), where we plotted all passenger scan D-minus

measurements with their respective OCP-ZY20 targets (X-axis scale removed due to confidentiality). Note that the difference in bar heights between the turnaround types is because we have a different number of observations for the two turnaround types, as flights depart more often from the platform.

Activity	Target	Actual average D-minus (start)	Standard deviation D-minus (start)
Departure Bus			
D-gate E90	D-27	46.76	19.44
Other	D-24	52.30	16.04
Baggage handling*	A+3	3.17	6.51
Catering*			
35 min E90	D-32	32.21	8.18
Other E90	D-48	57.58	22.18
35 min F70	D-31	30.02	8.68
Other F70	D-39	56.71	23.89
Cleaning*			
35 min Platform	A+9	13.39	5.45
Other Platform	A+9	31.06	22.40
35 min Semi-contact	A+7	11.27	4.02
Other Semi-contact	A+7	23.40	17.47
Fuelling			
35 min	A+0	10.59	8.11
Other	D-40	57.40	26.74
Passenger Scans			
Semi-Contact	D-17	20.68	5.53
Platform	D-24	25.62	4.50
D-gate E90	D-27	27.71	4.72
Toilet			
35 min F70	A+10	16.28	9.09
Other	A+10	20.12	74.11
Water*	A+1	20.29	21.43

Table 13 – Activity start performance

4.1.3 Activity starting time

Finally, we look at whether the activities started in time. Again, this is done relative to the OCP-ZY20 template. Note that some activities have a time window in the OCP-ZY20 that is larger than the required activity duration. These activities might thus start later than the planned time in the OCP-ZY20 without any negative consequence. Because of this, we expect some activities to appear to be underperforming based on starting time. For the activity starting time, we choose either to use the Dminus (start) variable or the Aplus variable, based on whether the target of an activity is arrival oriented or departure oriented. For a graphical representation of these variables, see Figure 7.

In Table 13, we left out data from the first of day turnarounds. This is done because of the format of the available data. Some activities have to perform two tasks, one in the evening and one in the morning. Cleaning of the aircraft for example happens in the evening, but sometimes the aircraft is serviced by maintenance during the night and thus the cabin has to be checked for any dirt or trash left by the mechanics in the morning. Currently, we can only see the start of the evening task

UNIVERSITEIT TWENTE.

and not that of the morning task. This skews the results greatly, with baggage handling being the largest, on average starting 1210 minutes before departure. Tasks where the first of day turnarounds are left out are marked with an asterisk.

Looking further at other activities and regular turnarounds, we see the same pattern as with the activity durations and completion times, good average performance with little to no departure-oriented averages being below target. On the other hand, as expected, we see many activities with an arrival oriented target time to be seemingly underperforming. For example cleaning, during a 35 minute turnaround at a semi-contact gate, its target starting time is 7 minutes after arrival. However, its average starting time is 11.27 minutes after arrival. This is due its available time window being larger than the required duration; it can thus start later but still finish in time.

Looking at the average activity duration, completion time and starting time, we conclude that most activities generally perform well when compared to their targets and no activities underperform structurally. However, standard deviation of each of these measurements show that there is still a large spread and many activities thus often finish after their deadline. In Section 4.2, we investigate whether correlations can be found between performance variables and departure delay and if such correlations exist, which performance variables correlate most.

4.2 Data Mining & Decision tree learning

In this section, we introduce decision tree learning. We discuss its background, the way it works and introduce the various parameters that we can set to control its behaviour. People already familiar with data mining, decision tree learning and random forests can skip to Section 4.3, where we apply these techniques to our data. This section is divided into six sub sections, each describing a part of the theory. First, we introduce machine learning (Section 4.2.1), after which we zoom in on decision tree learning in general (Section 4.2.2). In Section 4.2.3, we discuss how a decision tree is built, and in Section 4.2.4, we present overfitting, which is a serious risk when applying decision tree learning. We then continue in Section 4.2.5 by introducing feature selection, a way to improve decision tree learning when using data sets with a large number of variables. Finally, in 4.2.6, we discuss how missing portions of input data can be handled.

4.2.1 Machine learning

We start by giving a brief introduction into the technique decision tree learning and its background. The technique Decision tree learning is part of Machine Learning, which in turn is part of Data mining, an area of expertise that tries to identify useful patterns in large amounts of available data in an automatic or semi-automatic process (Witten & Frank, 2000). Machine Learning is one of the most important techniques used within Data mining to identify these patterns. Machine learning is a technique for finding and describing structural patterns in data, help explaining that data or make predictions from it. Input for a machine learning process is historical data; its output is a classification, applicable to both the historical data as well as new, unseen data.

There are two types of machine learning: supervised learning and unsupervised learning. Supervised learning divides the input data in predictor variables and a response variable and tries to classify the predictor variables based on the corresponding response variable value. If we look at the data presented in Table 14, we see some data taken from a health care dataset. For each entry (person), we have information whether he or she was a smoker or a drinker, his weight and the age at which he died (at a young age or at an old age). A good example of supervised learning would be to predict the age at which people die, based on their lifestyle. Thus, we would use the age of death as a response variable and the other variables as predictor variables.

The other type of machine learning is unsupervised learning. Here, we do not have a 'supervising' variable to which we relate the other variables, but rather every variable is equal. Thus, unsupervised learning tries to cluster its input variables together in any way possible, not just to a single response variable. Again looking at Table 14, unsupervised learning will look for any combination of these variables. For example, it might identify that if people smoke, they also tend to be a drinker. (Van der Aalst, 2011)

Smoker	Drinker	Weight	Age of death
Yes	Yes	110	Young
No	No	60	Old
Yes	Yes	100	Young
No	No	95	Old
No	No	70	Old
…	…	…	…
No	Yes	120	Young
Yes	No	80	Old

Table 14 – Example (sub)set of health care data (Van der Aalst, 2011)

4.2.2 Decision tree learning

For our goal, we clearly need a supervised learning technique, as we want to classify variables on whether they cause delay or not. The most used supervised learning technique is Decision tree learning. This technique builds a tree of predictor variables, with every branch being a certain value of a variable. A simple example is given in Figure 9, where we present a tree based on the entire dataset of Table 14. In this tree, we try to predict the age at which people die, based on the various variables present in the dataset. The way decision tree learning works is that it takes the entire dataset and tries to divide this dataset on every available variable. In our example, it will try to divide our data on the smoking variable, the drinking variable and the weight variable. Its goal is to gain a better classification of the response variable, where its ideal output would be a perfect classification, where in every end node all entries have the response variable value in which it was classified. Before we discuss the way the tree is 'learned', we give a small explanation on how to read a decision tree: We start with the root node, where we split the data into two branches, based on the first chosen variable, in this case 'smoking'. Each branch is classified, or predicted, to die at either an old or a young age. In this case, if the subject was a smoker, the person is
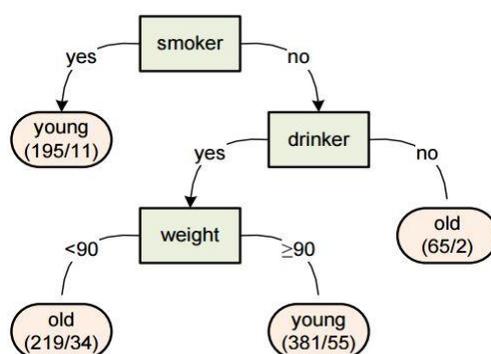
Figure 9 - Decision tree (Van der Aalst, 2011)

predicted to die young. If the subject was no smoker, we need more information to give a reliable prediction. We therefore look at whether the person was a drinker, if the answer is 'no', the person is predicted to die old. If the answer was 'yes', we create another branch on whether the person weighted over 90 kilo's or not. If this is the case, we predict him or her to die young; otherwise, the

prediction is that he will die old. Each variable split (square) is called a node, each group resulting from the split (line) is called a branch and a classification (oval) is called a leaf. All trees start with a single node and end with only leaves. The amount of nodes and branches in between can vary greatly and is not restricted.

4.2.3 Building a decision tree based on entropy

Now we know how to read a decision tree, we discuss how to decide which variables should be branches in the tree and at what place they should be placed. As mentioned earlier, the goal of a decision tree is to classify, or predict, the response variable as precisely as we can, using the predictor variables we have. This precision can be measured using the amount of variation in a single node, called entropy. Each step in a decision tree learning heuristic is designed to improve the total entropy of the model. For example, when we have not branched our dataset, we have a single group with a single prediction. This group has very high entropy, as all response values are in the same group, not just the ones for which we predicted their outcome correctly. When we split the dataset into two groups, based on a single property, we gain two groups as illustrated in Figure 10. We can clearly see that on average, the 'predictive classification' within the groups increased significantly, as the group that did smoke actually fits its prediction perfectly, while in the non-smoking group we only have single misclassification. We can expand this and further divide our data onto the property drinker, to gain even better prediction and thus lower entropy.
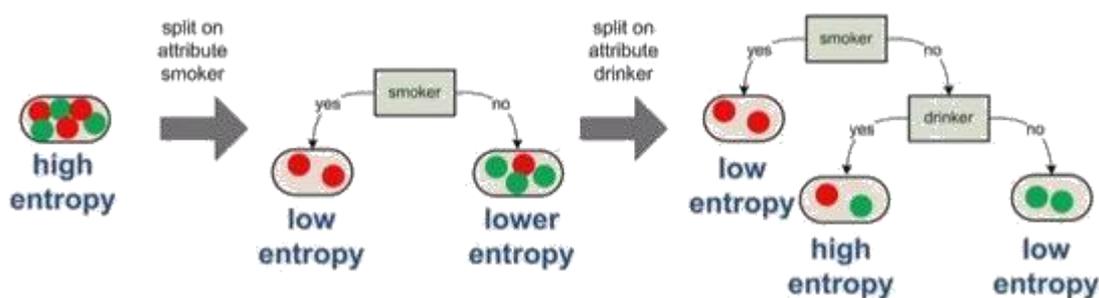
Figure 10 - Entropy gain (Van der Aalst, 2011)

At every step in a decision tree learning heuristic, the entropy gain of every available branching option is evaluated and the branch with the highest gain is chosen, this is called branching on information gain. Entropy can be calculated using formula (1), where k equals the possible values (in Figure 10 these are 'Yes' and 'No'). $p_i = \frac{c_i}{n}$ is the fraction of elements having value i (where i is each individual option in k) and $c_i \geq 1$ being the number of observations with value i and $n = \sum_{i=1}^{k} c_i$ being the number of observations that is predicted correctly in this leaf. In our case, k equals a maximum of two (young or old), entropy for the first node in Figure 10 equals 1, as shown by formula (2).

$$E = \sum_{i=1}^{k} p_i \, \log_2(p_i) \tag{1}$$

$$E_{Not\ split} = -\left(\frac{3}{6}\log_2\left(\frac{3}{6}\right) + \frac{3}{6}\log_2\left(\frac{3}{6}\right)\right) = -\left(\frac{1}{2} * -1 + \frac{1}{2} * -1\right) = 1 \tag{2}$$

If we split the data on the variable smoker, we can calculate entropy again by using the same formula. The 'smoker' leaf has entropy of 0, as shown in formula (3), the 'non-smoker' leaf has entropy of 0.811, as shown in formula (4). Total entropy of the tree is thus 0.54, as shown in formula (5). The total information gain if branching on the variable smoker is thus 1 – 0.54 = 0.46.

UNIVERSITEIT TWENTE.

KLM

By calculating the potential information gain from branching on all available variables, decision tree learning will determine the branching option with the highest information gain and apply it to the tree. If this process is repeated until all leafs cannot be branched again, or the projected information gain is negative, the decision tree learning process will stop and present the final tree. (Van der Aalst, 2011)

$$E_{Smoker} = -\left(\frac{2}{2}\log_2\left(\frac{2}{2}\right)\right) = -(1 * 0) = 0 \tag{3}$$

$$E_{Non\ smoker} = -\left(\frac{1}{4}\log_2\left(\frac{1}{4}\right) + \frac{3}{4}\log_2\left(\frac{3}{4}\right)\right) = -\left(\frac{1}{4} * -2 + \frac{3}{4} * -0.415\right) = 0.811 \tag{4}$$

$$E_{Total\ tree} = \left(\frac{2}{6} * 0 + \frac{4}{6} * 0.811\right) = 0.54 \tag{5}$$

The method described above will work fine for Boolean or categorical data. However, with numerical data, such as the variable weight in our dataset, the decision tree learning process does not have a fixed value on which it can branch. An algorithm called "C4.5" provides a solution to this problem. We do not discuss its working in detail, but dependent on the size and variance of the input data, it evaluates various splits per variable to find the optimal value to split the data. If the number of unique values of this variable is sufficiently small, it will try a split on every value. If this number is too big (for example a continuous variable), it will branch-and-bound various splits to approximate the optimal value (Witten & Frank, 2000). For example, in the decision tree in Figure 9, the value for the variable weight that provides the largest information gain was found to be 90 kilo. The C4.5 is, as is the case with nearly all decision tree learning implementations, used by our decision tree learning algorithm.

4.2.4 Overfitting

Decision tree learning has an important trade-off. Imagine that we have an additional variable in our dataset, indicating whether subjects are professional chess players. However, in our dataset, we only have a single professional chess player, which happened to have died at a young age. The decision tree learning process will evaluate the variable chess player and find that all subjects that match this criterion (one) have died young, thus branching would improve entropy slightly. In reality however, it makes no sense to branch on this variable, as we would need more chess players in the data set to be sure this relation actually holds. We arrived in a situation called overfitting, where decision tree learning has created too many branches, which do not necessarily improve the overall classification. A decision tree learning process accepts a few parameters to prevent overfitting; these can be found in Table 15. The first is the minimum number of data entries that should be present in a node if the algorithm tries to branch it. For example, if this is set to ten, the algorithm will ignore any nodes that have nine entries or less, thus making that node a leaf. The second is the minimum number of data entries required in a resulting leaf. Our chess player example would not be accepted if this parameter were set to be two or higher, as one of the resulting leaves has a single data entry. The third parameter is the minimum information gain per branch. Our calculated example of 0.46 information gain would only be accepted if this parameter is set to 0.46 or lower. The final parameter, the maximum tree depth, prevents our tree from growing very large and thus complex as it limits the size of the largest sequence of nodes within the tree. A long tree reduces readability and usability.

These parameters prove very useful in large amounts of data with many variables to branch on, as in our own case, we have roughly 47,000 turnarounds to analyse and we are not interested in leafs with only one or two turnarounds in them. However, these parameters should also be used with caution, as setting them too strict means that we might lose valid classifications that do add value. (Witten & Frank, 2000)

Description	Example value
Minimum number of entries in a node before branching that node	10
Minimum size of each resulting leaf needed for accepting a branch	3
Minimum information gain per branch	0.1
Maximum depth of tree	30

Table 15 – Decision tree parameters preventing overfitting

4.2.5 Feature selection and the Random forest technique

Another risk in decision tree learning is that the algorithm has too many variables to choose from. When many variables are present, there is a chance that the tree will branch on variables that show a good information gain early in the tree, but result in little to no information gain later on. While, if the tree would have accepted less information gain in the beginning, information gain further down the tree would be much better. Feature selection is a technique that can be applied to the data before building the tree to prevent this from happening. There are various ways to execute feature selection, all revolve around reducing the amount of available variables while still keeping most of the classification strength of the overall data. (Mladenic, 2005)

The method we use for feature selection in this research is based on the technique Random Forest. This technique creates a large number of randomly selected subsets of the available variables. Each subset is randomly selected from the entire set of variables, so (when a large number of subsets is created) each variable is present in many different subsets. It then continues to apply decision tree learning to each of these subsets, thus creating many trees, each of a different shape due to the different variables present. Every tree is then ranked based on its classification power, that is, the percentage of correctly predicted subjects when the tree is applied to its own training dataset. The relevance of each variable is measured by the performance of all the trees that contain that specific variable. For example, when a specific variable is present in three trees, the performance of that variable is the average performance of these three trees. By making a large number of subsets (trees), we make sure that many combinations of variables are tested, to gain a good average performance of each variable and prevent that a single combination of two variables gain a dominant position in our tree (Reif, Motsinger, McKinney, Crowe Jr, & Moore, 2006). Table 16 lists the input parameters needed for the random forest technique.

Description	Example value
Number of trees to build	5000
Minimum leaf size needed for accepting a branch	10
Number of variables to draw for each tree	10

Table 16 – Random forest input parameters

4.2.6 Missing data handling within the random forest

Before applying decision tree learning on our dataset, we still have a few difficulties to overcome. First, for some variables we have a large number of missing values. To name a few examples, we use data from the ACARS system to verify when the various doors of the aircraft were opened and closed. However, the Fokker 70 aircraft, which is nearly half of the KLC fleet, does not support this ACARS system, thus reducing the availability of this data by half. Also, because some turnarounds

are performed at the Semi-contact B-pier at Schiphol, not all turnarounds have passenger bus data present. Extra fuel is only required by some of the flights, making the data from all extra fuel related variables mostly absent. Because we filtered our dataset in Section 3.4.1 to exclude turnarounds where important variables are missing, all empty values in our dataset represent activities that were not needed that specific turnaround, and thus actually not missing, but simply not present.

However, random forest creation is a robust technique that can handle some missing values, but the amount missing from our data proved too much, resulting in strange classifications and runtime errors. Therefore, in columns with a large number of empty values, we had to fill these empty positions with values that do not significantly influence its prediction but allow the random forest algorithm to process these variables. There are many techniques to fill missing values, each with their own advantages and disadvantages. Simply replacing the missing values with the average of the other values, which used to be a popular technique, has been discredited. Researchers claim it may produce biased estimates of the regression coefficients (Allison, 2009). Other, more advanced, techniques are often based on estimating the missing values based on values of the other variables present (Graham, 2009). These techniques assume that the missing values should have had a value, but somehow this value got lost during measuring or processing of the data. However, because we filtered our dataset to exclude turnarounds with critical measurements missing, the large number of values that is missing is because they were simply not executed. Thus estimating a value from other variables would only lower our performance, as the decision tree learning process might predict delay with a filled-in value that was not even present during that turnaround.

Therefore, we use a simple approach to fill our missing data. For each variable type (e.g. D-minus, Aplus), we determine a 'good' value that, if the corresponding activity was actually performed, it would not have influenced departure delay, and place this value in the empty positions. For example, for the 'D-minus' variable of an activity (the time between activity completion and scheduled departure), this value was set to be 61 minutes. Given a minimum turnaround time of 35 minutes, an activity that is completed 61 minutes before scheduled departure will have had no influence on departure delay. Similar with lateness, when an activity is finished 61 minutes before its scheduled completion time, it will have no influence. The values have been chosen to represent rather extreme values, to be able to identify them from the regular values. Each data type for which we have defined missing value replacements can be found in Table 17.

Note that for the door data on the Fokker 70 flights, the assumption that missing values did not influence departure delay does not hold, as they are an important step in the turnaround process. However, because we split our data set to the various turnaround types and thus by aircraft type, each data set has either all door data present (Embraer turnaround types), or no door data at all (Fokker turnaround types). In the latter case, this results in the algorithm ignoring these variables entirely for these turnaround types, thus no missing value replacement is needed. Unfortunately, this reduces the value of the results for the Fokker turnarounds, but as KLC has announced to phase-out all Fokker 70 within three years, this compromise is accepted.

Data type	Dummy value
D-minus	61
A-plus	-61
Duration	-1
Lateness	-61
LateStart	-61

Table 17 – Replacement values for missing data

UNIVERSITEIT TWENTE.

4.3 Analysis into causes of delay

Which circumstances or activities actually have an impact on D0 performance?

In order to identify relations between circumstances and departure delay, we use two methods. First, we apply decision tree learning to look for predictor variables in our dataset that correspond to departure delay. For this purpose, we choose our experimental design in Section 4.3.1, where we determine the input parameters for our algorithm. Next, in Section 4.3.2, we present and discuss the results. We have visualised the steps of the decision tree learning process in Figure 11. In addition, we use the delay propagation method we introduced (Section 3.2.1) to look for activities within a turnaround that were delayed and, according to the propagation algorithm, delayed the entire turnaround.
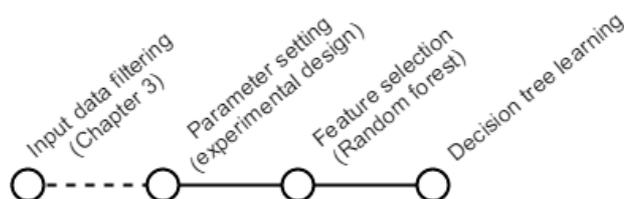

Figure 11 – Decision tree learning process

4.3.1 Experimental design

Before presenting the results, we first discuss how we instruct our decision tree learning process which data to use and what settings it should take into account. In other words, how we make sure that the results from our analysis are relevant and not influenced by selecting bad parameter settings. In order to select the correct parameters values, we used multiple runs, each with a different combination of parameter values.

Because of the large number of parameter values that we have to choose, we divide them into three categories, namely input data related, random forest related and decision tree learning related. When testing the different parameter values, we perform an incremental design. First, we test the various combinations for the input data related parameters. From these, we select the best predicting combination and continue to test the performance of the second set, the random forest related parameters. Again, this results in a set of best performing parameter values, which we then use to determine the best combination of parameter values in the decision tree learning related category. This will result in a good combination of parameter values that we can feed to our algorithm.

Selecting input data related parameter values

In Chapter 3, we presented the dataset we use as input for our decision tree learning process. However, within this dataset, we have a few options regarding what data we actually use, as we have various variables that describe similar performances. For example, for each activity, we have the time between scheduled departure and activity completion time, but also the deviation of activity completion time to its target completion time (Table 7). As these variables represent similar performance, we need only use one of these types of data. Next to that, we also want to test whether using the dummy data generated for the random forest (Table 17) would also improve performance for our decision tree learning process.

As mentioned in Section 4.2, a decision tree learning process accepts certain parameters with which we can control its behaviour. We already introduced the parameters preventing overfitting, but there is also a parameter regarding its behaviour around missing values. In Chapter 3, we filtered out all turnarounds with measurements of one or more critical activities missing. In Section 4.2.6, we introduced dummy values for activities that were not executed, for example

38

when water was multi stretched. However, we still have some variables where missing values are present, for example wind speed or wind direction. Because cases where these variables are missing represent a very small portion of our data set (< 1%), we do not fill them with replacement values or dummy data, but let the decision tree learning algorithm decide what to do with them.

When it finds missing values in a variable on which it wants to branch on, it can choose to drop the observation, thus not classifying it any further, or it can send the observation to the branch with the majority of observations. As this setting has a major influence on the performance of the process when testing to implement dummy data or not, we tested each dataset with each of these options. This resulted in eight unique configurations, which, together with their performances, can be found in Table 18.

Parameter:	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run8
Activity variables	Lateness, LateStart, Duration				Dminus, Aplus, Duration			
Missing values filled with dummy data	Yes		No		Yes		No	
Send missing observation with majority or drop	Keep	Drop	Keep	Drop	Keep	Drop	Keep	Drop
Performance	75.7%	75.7%	75.4%	72.5%	81.0%	**81.3%**	80.8%	77.5%

Table 18 – Decision tree learning input data

Looking at this data, we see that while the performance of runs 5, 6, and 7 all are within 0.5% of each other, run 6 is the most successful, with 81.3% of all values correctly predicted. This makes sense when looking at the different parameter settings. Run 6 is set to not classify turnarounds with a missing observation for the splitting variable any further, rather than to send them with the majority of data. Because the values of variables that are missing due to not being executed are already filled with dummy data in runs 5 and 6, the missing values in these runs might be relevant but got lost during measuring or processing. If we send turnarounds with that specific measurement missing with the majority of observations, we effectively classify them as being not delayed, as the majority of flights is not delayed. This means that all turnarounds with missing values of splitting variables that actually had an influence on departure delay are classified wrong in run 5 and 7. Because we only have a small number of missing observations in our data due to the data cleaning in Chapter 3, the gain from this parameter is small. However, it assures us that run 6 is indeed the best combination of settings.

Selecting random forest related parameter values

Now we have selected the input data and the parameter settings for how the decision tree should handle missing values, we try to find the optimal setting for our random forest input parameters. In Table 19, the four configurations that we test can be found. First, the number of trees must be determined. In 2012, Oshiro, Perez, and Baranauskas performed a study on the amount of trees that were needed to perform a good feature selection. They tested various data sets, each with a different size and complexity, and various forest sizes, ranging from 2 to 4096. They concluded that, with a large and complex data set, 4096 trees still showed an improvement over 2048 trees. Therefore, given the large number of variables in our dataset and an abundance of computational power, we choose to test two forest sizes, 1000 and 5000 trees. Even at 5000 trees, computation time is within reasonable limits. Run 11 and 12 from Table 19 each completed in roughly 30 minutes on a PC with a 3.9 GHz i7 CPU.

Next to the number of trees, we can control some properties of each single tree, just like with a regular decision tree learning process. The minimum leaf size, the number of observations that match a branch, should be large enough to prevent overfitting but small enough to allow the

UNIVERSITEIT TWENTE.

KLM

process to develop a sufficient number of branches. Given the large spread in the size of the different turnaround datasets with over 11,000 observations in the largest (F70, platform, standard duration) and 352 for the smallest (Embraer, D-gate, first of day), this trade-off is even more important. We test 2 sizes, 30 and 60, with the latter being actually too large for the smallest dataset. However, given the small number of occurrences and because the turnaround type shows strong similarities with other turnaround types (and thus their results might be applicable to this turnaround type), we accept this shortcoming if it improves the performance at larger datasets.

Finally, we set the number of variables that have to be included in each tree to be 10. Literature suggests (Witten & Frank, 2000) to include the square root of the total number of variables present. For example, when there are nine variables in the total set, randomly select three variables per tree. A higher number of variables per tree would reduce the number of trees needed but at the same time lessen the strength of the feature selection (as the resulting trees will be more similar). A lower number increases the number of combinations that can be made and thus improves feature selection, but as a trade-off requires more trees to be built to be reliable. In our datasets, we have between 120 and 160 variables, depending on the turnaround type (For example, Semi-contact turnarounds do not have bus related variables and Fokker 70 turnarounds do not have door data related variables). So setting 11-13 variables per tree would be recommended by literature. However, given the large number of trees that we test, we can lower this number, increasing the quality of the feature selection. An indication that 10 variables would be too little for our forest sizes would be that the runs with 1000 trees would perform significantly worse than the runs with 5000 trees. If these runs perform similar, 10 variables per tree is large enough to be reliable while at the same time (because it is smaller than the square root of the total number of variables) small enough to ensure a proper feature selection.

Parameter:	Run 9	Run 10	Run 11	Run 12
Number of Trees to build	1000		5000	
Minimum leaf size before accepting branch	30	60	30	60
Number of variables for each tree	10	10	10	10
OOB rate	21.8%	19.5%	**19.0%**	19.4%
Performance of decision tree created from selected variables	81.3%	81.3%	**81.3%**	81,3%

Table 19 – Tested random forest input parameters

In Table 19, we listed the various runs and their performance. The OOB rate (or Out Of Bag rate) is the average number of observations that could not be classified with the variables by the various trees in the forest. A lower OOB percentage indicates a higher quality of predictor variables, which is preferable. Next to the OOB rate, we list the prediction performance of the decision tree build from the selected variables. While OOB rates differ among the different runs, the decision trees perform identical. This means that each random forest configuration is able to select enough good variables from which the decision tree learning process is able to fit an identical tree. Therefore, each set of random forest parameters we tested provides a similar quality of selected variables. To improve the efficiency of our process and reduce calculation time, we could test additional sets of parameters (for example a smaller number of trees) to see with which settings the random forest would start to not be able to provide a high enough quality of variables to build a good decision tree. However, because the calculations for this research are non-recurring, there is no necessity to do this. The calculations needed to determine an optimal setting would offset any gains.

Selecting decision tree learning related parameters

UNIVERSITEIT TWENTE.

KLM

Lastly, we have to set the parameters for our decision tree learning process itself. As discussed earlier, we have four parameters at our disposal, presented in Table 15. We test these with the sixteen different combinations that are listed in Table 20. The first three settings are similar to parameters we have tested before and are self-explanatory. The fourth one, maximum depth, determines the maximum depth of the tree. We test this parameter for two settings; 30 and 5. This means that the decision tree may be no more than 5 and 30 branches deep respectively. A deeper tree might result in a better prediction, but there is the risk of overfitting and readability / acceptance by the business is also severely limited when a tree is 30 branches deep. We do not want to account for very specific circumstances in which 30 variables are of influence, but rather check for circumstances that are more general. We therefore also test for a maximum of 5 branches. Depending on the performance of this tree compared to the deeper tree, we might opt to select this tree over the more complex tree because the individual variables have more predictive power and usability is improved.

Parameter:	Run 13	Run 14	Run 15	Run 16	Run 17	Run 18	Run 19	Run 20
Minimum size of leaf before trying branch	60				120			
Minimum size of leaf before accepting branch	30				60			
Minimum information gain before accepting branch	0.005		0.05		0.005		0.05	
Maximum Depth	30	5	30	5	30	5	30	5
Performance	**81.3%**	79.6%	78.2%	78.2%	79.1%	78.5%	77.4%	77.4%

Parameter:	Run 21	Run 22	Run 23	Run 24	Run 25	Run 26	Run 27	Run 28
Minimum size of leaf before trying branch	120				180			
Minimum size of leaf before accepting branch	30				60			
Minimum information gain before accepting branch	0.005		0.05		0.005		0.05	
Maximum Depth	30	5	30	5	30	5	30	5
Performance	80.5%	79.5%	78.2%	78.2%	78.9%	78.3%	77.4%	77.4%

Table 20 – Decision tree learning process input parameters

The best performing configuration proved to be run 13. This means that with the settings from this run, the decision tree learning process was able to select the best variables on which to branch on in our datasets. We therefore use the results from run 13 for our analysis. To summarize our experimental design, we listed the input and parameters we used for our analysis in Table 21.

Parameter	Value
Input data related:	
Data division	Turnaround type
Selected activity related variables	Dminus, Aplus, Duration
Missing values filled with dummy data	Yes
Missing observations with majority or drop	Drop
Random forest related:	

UNIVERSITEIT TWENTE.

KLM

Number of trees:	5000
Table continues on next page	
Minimum resulting leaf from branch	30
Number of variables per tree	10
Decision tree learning related:	
Minimum size of leaf before trying to branch	60
Minimum resulting leaf from branch	30
Minimum information gain from branch	0.005
Maximum depth of tree	30

Table 21 - Selected experimental design

4.3.2 Decision tree learning results

We applied a decision tree learning process with the settings from Table 21 to sixteen datasets, one for each turnaround type and one large dataset with all turnarounds included. As an example, the tree resulting from the 'Fokker 70, QTA, Semi-Contact' is illustrated in Figure 12. The performance of the different trees in each dataset can be found in Table 22. It is interesting to see that the Embraer datasets produce significantly better performing trees, most likely due to the presence of door data, which provide opportunities to predict delay. The full dataset, with all turnaround types included, proved to be relatively hard to predict correctly. This indicates that some turnaround types indeed require different variables, and the turnaround types differ significant enough to justify dividing them. However, many trees still show strong similarities and because of this, and to promote readability, we do not discuss every tree in detail, but rather discuss the results in general. This means we discuss which variables are most commonly present in the different trees. Note that due to confidentiality the numbers in this section are altered and cannot be used as a reference for actual performance.
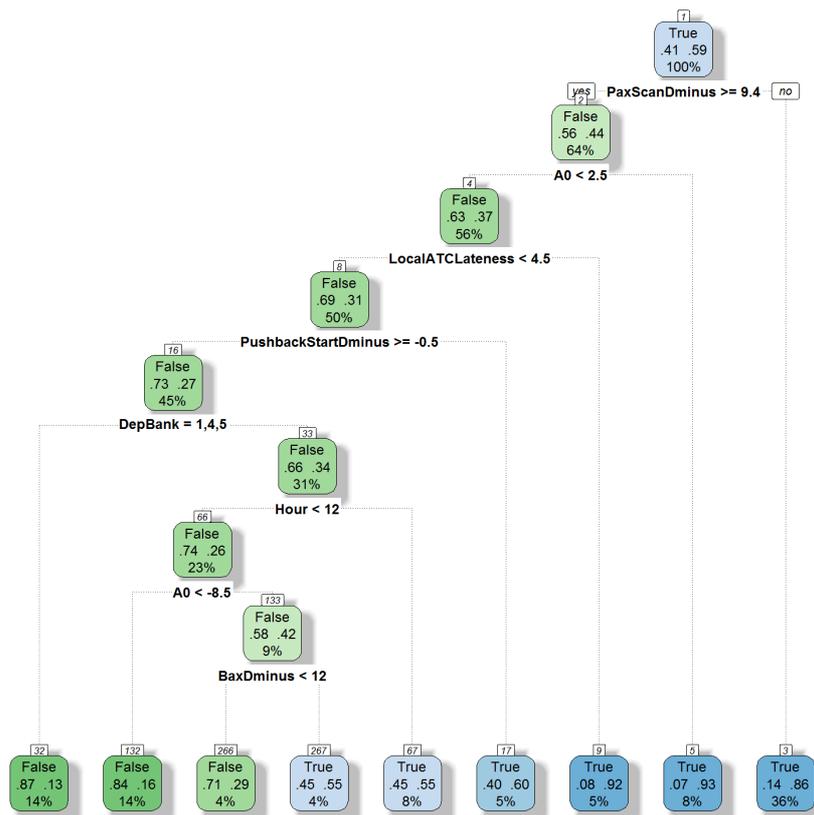

Figure 12 - Example decision tree belonging to Fokker 70, QTA Semi-contact turnaround

	1	2	3	4	5	6	7	8	9
Aircraft type	Embraer E190								
QTA	Yes	No		Yes	No		Yes	No	
Gate type	Platform			Gate D6			Semi-contact		
First of day	No		Yes	No		Yes	No		Yes
% Correctly predicted	86.7%	84.9%	82.9%	87.8%	84.9%	83.5%	87.2%	81.3%	82.5%

	10	11	12	13	14	15	16
Aircraft type	Fokker F70						All
QTA	Yes	No		Yes	No		All
Gate type	Platform			Semi-contact			All
First of day	No		Yes	No		Yes	All
% Correctly predicted	76.3%	72.0%	74.7%	80.8%	78.2%	75.6%	77.8%

Table 22 – Performance of each tree

In Table 23, we list the most important variables, both in general and for specific circumstances. Of course, the specific predictive power of the variables varies among the different turnaround types, but it still gives a good general overview. The most common splitting value represents the value where the observations are divided to be predicted either delayed or not delayed. When the stated condition is met (for example "<5.5" for Local ATC lateness), it means that when a turnaround has a value for Local ATC lateness of 3 minutes or smaller, the turnaround is predicted to be not delayed. Of course, for variables deeper down the tree, the considered values are filtered by preceding branches, thus the suggested splitting value is only valid for that subset and may not be generalised over the entire dataset. However, some variables are at different depths within the different decision trees, all with a similar splitting value, indicating the splitting value might hold for most of the turnarounds within each dataset.

Variable name	Most common splitting value
Most common variables:	
Local ATC lateness	< 5.5
PaxScan Dminus	> 8
A0	< 0 for QTAs, < 10-15 for non-QTA
Pushback Start Dminus	< 0.5
Load factor	< 91%
Specific E190 variables:	
Ready to depart lateness	< 3
Specific Platform variables:	
DepartureBus Dminus	> 3.5

Table 23 – Most common variables with their respective splitting values

The most important variable proved to be Local ATC lateness. In 14 out of the 15 turnaround types, it was the first or second variable in a tree, each time with a splitting value of roughly 5.5 and a correct classification of (when predicting to be delayed) roughly 90%. This means that in 90%

UNIVERSITEIT TWENTE.

of the cases when there are more than 5 minutes between the ready-to-depart notification and clearance from local ATC, the flight is delayed. This makes sense, as a ready-to-depart should only be given when the flight is fully ready for departure, and the turnaround schedule is designed to have this moment at the scheduled moment of departure. Thus, by design, any Local ATC lateness results in a delayed flight.

The second most important variable, also present in 14 out of 15 turnaround types, is PaxScan Dminus (the time between scanning of the last boarding pass and scheduled departure). In nearly all turnaround types, this should be more than 8 minutes, or the flight is predicted to be delayed. These 8 minutes make sense, as there are many activities that have to take place between scanning the boarding passes and departure. The passenger has to walk down the gate, in case of a platform departure be transported to the aircraft, board the aircraft and take his seat. At the same time, the load sheet has to be printed and carried aboard, the passenger door has to be closed, stairs removed, and an optical check of the aircraft has to be performed. Important to note is that the splitting value for Semi-contact departures ranges between 7 and 8 minutes, the splitting value for the platform and D6-gate departures ranges between 8 and 12 minutes. This seems to correspond with the KLC procedure deadlines of 8 and 9 minutes, with the latter being fairly optimistic when compared to the actual data.

Next, A0, or arrival punctuality, is also an important variable. For the QTA turnarounds, it should be negative to prevent delay. This means that the incoming flight cannot be delayed, or the departing flight will also be delayed. Again, this makes sense, as in a QTA the planned ground time is equal to the minimal time that is needed to perform all activities. This means that, when the arriving flight is delayed, the ground process has little to no slack to make up for it, and will simply propagate the delay to the departing flight. For longer turnarounds, the A0 is, logically, of less influence. In most non-QTA Embraer 190 decision trees, A0 is absent, meaning the decision tree learning algorithm was not able to find a significant correlation between A0 and departure delay. On the other hand, for Fokker 70 turnarounds departing from the platform or semi-contact, the A0 should be less than 12 or 16 minutes respectively. This means that an incoming flight can be delayed for up to 12 minutes without direct consequences for the departing flight. This difference can be explained, as the standard turnaround time of an Embraer is 50 minutes, whereas for the Fokker 70, this is 40 minutes. That means that an Embraer turnaround has more slack, as the minimum ground time for both types is 35 minutes.

Another important variable (present in 8 out of 15 turnaround types) is PushBackStartDminus, the time between the starting moment of the pushback truck operator and scheduled time of departure. Again, this correlation can be explained rationally, because the splitting value was always equal to 0.5. This means, that when the pushback truck operator started its task after scheduled departure, the flight is delayed, which is obvious. The presence of this variable in our trees suggests that the pushback truck might not always be on time at the aircraft to start in time, thus delaying the flight. However, there is some uncertainty on the reliability of this data. The pushback truck operator has to perform multiple actions before he can start actually pushing the aircraft. First, he has to position himself in front of the aircraft and link his truck to the landing gear. Then, when he received confirmation from the pilot, he will start pushing the aircraft. Time between linking and actual pushing can be up to 20 minutes in extreme cases, when the pushback truck operator is in time, but the aircraft is delayed for other reasons. The problem with our data is, that some operators start their task in CHIP when they start linking to the aircraft, whereas other operators start their task when they start the actual pushing. In the first case, a 'pushbackStartDminus' of 0 minutes or less indicates that the pushback operator was the cause of the delay, in the latter, this might not be the case, as he may have been already ready for 5 or 10 minutes.

The final general variable is load factor. Regardless of aircraft type, when load factor is above 91%, the flight is predicted to be delayed. A logical explanation is that the duration of some activities depends on the number of passengers present. For example, the number of bags to be placed into the cargo hold increases, the time needed to scan all boarding passes increases, the time needed to get everybody seated within the aircraft and cabin luggage stowed away increases, and so on. Target processing times were determined based on activity durations under full load, but without disturbances. Thus, a single disturbance can disrupt the process, lengthen activity duration and delay the flight, because there is no slack time to make up for it. When the load factor is low, activities have more slack, allowing for some disturbances.

Finally, a variable that was only present in Embraer turnarounds: the RTD lateness. This is the time between the closing of the last door and the captain notifying the tower that he is ready to depart. RTD lateness might be caused when passengers have not taken their seat yet, or cabin luggage that has not been stowed away properly. Especially for flights with a high load factor this is an issue, as the amount of luggage people carry will be close to the maximum capacity of the luggage holds, and people will have to search for a place to stow their luggage. Nevertheless, although we could not determine the RTD lateness for the other turnarounds, because we do not have door closure data from the Fokker, we expect the RTD lateness to have an influence on Fokker turnarounds as well, because luggage and seating issues also occur in the Fokker.

4.3.3 Propagated and contributed delay

In this section, we compare our results with the delay causes suggested by the delay propagation method used in the AP2K project. The calculations needed in this section were done by KLM internally. For this research we simply present the results and compare them to our own results. For fair comparison purposes, we only use data from turnarounds that are present in the data set gained in Section 3.4. The contributed delay of each event is determined using a two-step calculation. First, using the delay propagation method (Section 2.3), the propagated and contributed delay of each turnaround are determined on a schedule level. This means that all flights of a single aircraft on a single day are linked, and the propagated and contributed delay of each flight and turnaround are calculated. Second, the algorithm tries to distribute the contributed delay of each turnaround within this turnaround. For example, when a turnaround has caused 10 minutes of contributed delay, it will distribute this delay over the various tasks by using the OCP-ZY20 planned times, dependencies and actual performance. The result might for example be that, of these 10 minutes, 5 minutes were caused by the late closing of the cargo doors, while 5 minutes were caused by a late start in boarding pass scanning. Note that due to confidentiality the numbers in this section are altered and cannot be used as a reference for actual performance.
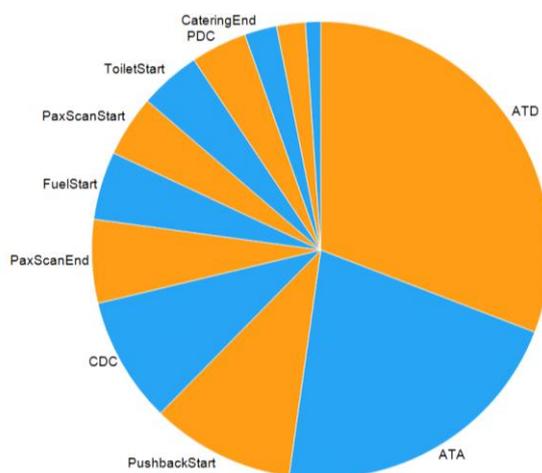

Figure 13 - Event contributed delay, summed over all turnarounds

45

UNIVERSITEIT TWENTE.

To get an overview of which activities or events cause the most delay according to the contributed delay algorithm, the 'event contributed delay' resulting from the previous steps is summed over all turnarounds in our data set. In Figure 13, the events that caused at least 1% of the total delay are visualised. Note that for the AP2K project, the ready to depart notification is not implemented in the algorithm yet, so no distinction can be made between Local ATC delay and ready to depart delay. These are combined into a single event: actual time of departure (ATD). Looking at Figure 13, we conclude that this combined ATD delay is by far the most important cause. This is in line with our findings from the decision tree learning process, where we found that Local ATC delay and ready to depart delay are important variables in predicting delay. Other variables also match with our previous findings, namely the cargo door, passenger door, pushback truck and passenger scans. This indicates that, while they are very different ways of determining delay causes, both algorithms show similar results.

However, there are differences. For our decision tree learning algorithm, we used additional data sources that the delay propagation algorithm did not have. Also, the delay propagation algorithm can only look at timed events, rather than also use variables that cannot be placed on a time line such as aircraft position or weather conditions. For example, whereas the decision tree learning algorithm tells us a high load factor is a cause of delay, the delay propagation algorithm cannot provide such insights. To get a better comparison, we assigned the propagated delay from each flight as contributed delay of the actual time of arrival (ATA). This way we see that a large portion of delay is propagated towards the arriving flight, this is in line with the large influence of arrival punctuality we found with our decision tree learning algorithm.

Looking at the results of both the decision tree learning algorithm and the delay propagation algorithm, we see similar results. Given the different nature of the two algorithms, they can be used in a complementary fashion. The delay propagation method gives a quantitative insight into which tasks or events contribute most to departure delay. It can make a distinction between tasks that were delayed because of preceding tasks or events, and tasks that were delayed on their own account. Decision tree learning provides a good overview of which variables correlate with departure delay, these can either be timing related variables, but also circumstances. It provides threshold values for each variable, indicating for which values delay is predicted.

Event	Difference of unfiltered data to filtered data
ATD	- 0.4%
ATA	+0.2%
PushbackStart	- 0.1%
CDC	- 0.1%
PaxScanEnd	+0.5%
ToiletStart	- 0.2%
PDC	- 0.1%
PaxScanStart	- 0.1%
FuelStart	+0.0%
CateringEnd	+0.0%

Table 24 – Difference in percentage of total delay, filtered data versus unfiltered data

Finally, we did a check on the delay propagation data provided by the AP2K project. For comparison purposes, the data in Figure 13 is limited to include only the filtered turnarounds that we used in the decision tree learning algorithm. However, in Chapter 3, we discarded roughly one third of our data due to data errors and missing data. Because the contributed and propagated

UNIVERSITEIT TWENTE.

KLM

delay is already calculated for the entire data set for the AP2K project, we can easily compare the results from the entire set with the results of the filtered data set. This comparison is presented in Table 24. Although small differences occur, no major differences exist between the two sets. Of course, this is just a very rough comparison and further research is required before any hard conclusions can be drawn, but it seems that the filtered data is a good representation of the entire data set.

4.4   Reliability of the current delay identification method
Based on our previous findings, is the current delay identification method reliable?

In this section, we discuss the quality of the current delay identification method by comparing the delay attributed to a specific code to our data. Because it is the most frequently assigned delay code, (over 18% of all flights receive a code 89) and we have good comparison values in our data, we perform this comparison primarily on code 89, which stands for 'Restrictions at departure airport', or the most common cause: no departure slot made available by local air traffic control (Local ATC). This code should be assigned when the aircraft is fully ready to depart and departure is requested at local ATC (RTD notification), but local ATC does not have a free departure slot available, ordering the aircraft to wait at the gate before it can be pushed back. Because a TLO cannot see whether an aircraft is not departing due to an unready cabin or because of Local ATC restrictions, KLC suspects code 89 to be given too often.

Using our data, we can determine the actual delay caused by Local ATC by calculating the time between the Ready to depart notification and actual departure, which we call Local ATC lateness. When the delay code has been assigned correctly, this amount of minutes should be equal to the amount of delay minutes assigned to code 89, either as primary delay code or as secondary delay code. In Figure 14, we plotted the comparison between these two, where we allowed a minute of

Comparision between 'LocalATCLateness' and Assigned Code 89 minutes (OK = max 1 minute deviation)
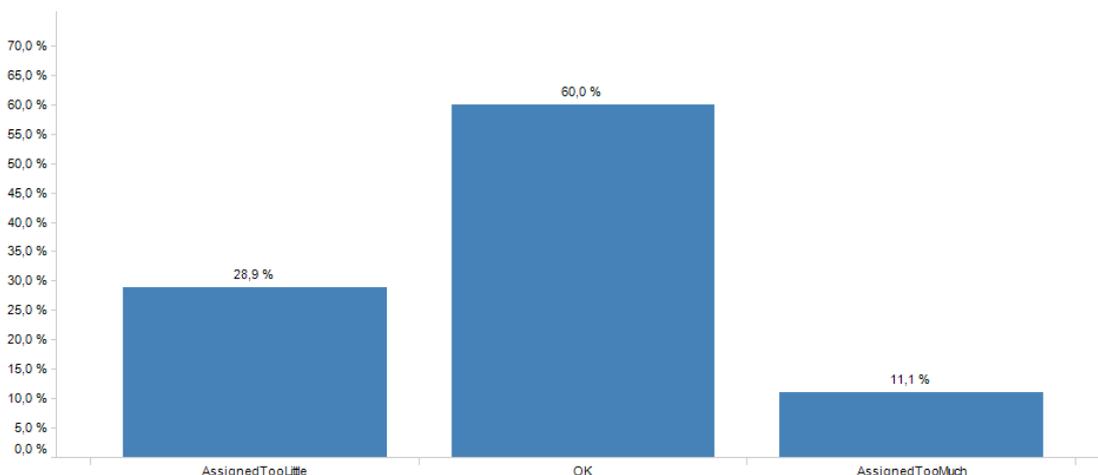
Figure 14 – Code 89 minutes compared to the actual ATC delay

deviation between the two times due to rounding errors or human error.

First thing to note is that for 60% of all flights, code 89 assignment seems to be correct. Note that this includes flights where no ATC delay occurred (and none was assigned) and flights that did have an ATC delay, but these were attributed correctly. Next, for 11% of the flights, at least two minutes of delay were attributed too much to Local ATC by means of code 89. This seems to match with KLC suspicions that the code is being used too often. These two situations are graphically represented in Figure 15, where 'actual ATC delay' indicates what the actual ATC delay was in this scenario, 'assigned ATC lateness' indicates the amount of ATC delay that our calculation

UNIVERSITEIT TWENTE.

determined, and 'assigned code 89 minutes' indicates the amount of ATC delay that was assigned by the TLO.

However, in Figure 14, we also see that in roughly 30% of the cases our data suggests that the amount of assigned Local ATC delay is actually too low when compared to the time between the ready to depart notification and actual departure. Looking deeper in our data, we notice that for over 400 flights the RTD lateness (time between all doors closed and RTD notification) was negative, indicating that the RTD notification was given before the doors were closed, which would be impossible by its definition. KLC Flight Operations confirmed that this indeed is a possibility, as sometimes pilots give a RTD notification prematurely when they expect a long Local ATC delay. For example, the pilots expect they will be ready for take-off in 5 minutes but they notice on the radio that the current average waiting time for a departure slot is 10 minutes. In such a situation, they might give the RTD notification prematurely, to prevent having to wait the full 10 minutes after they are ready. Although the local ATC is indeed causing delay in such cases, it cannot be held responsible for causing 10 minutes of delay (the time between RTD and actual departure) which we currently do in our comparison, but only for the time after the aircraft is actually ready. In some cases, pilots may expect a delay in slot assignment, where in reality there proves to be none. In this case, the pilots give the RTD notification 5 minutes early, after which they are granted a departure slot immediately. At this moment in time, the aircraft still has to be prepared, still leaving 5 minutes after the RTD notification. In our comparison, these 5 minutes would wrongfully be attributed to Local ATC. A graphical explanation of this is represented in the left figure of Figure 16.
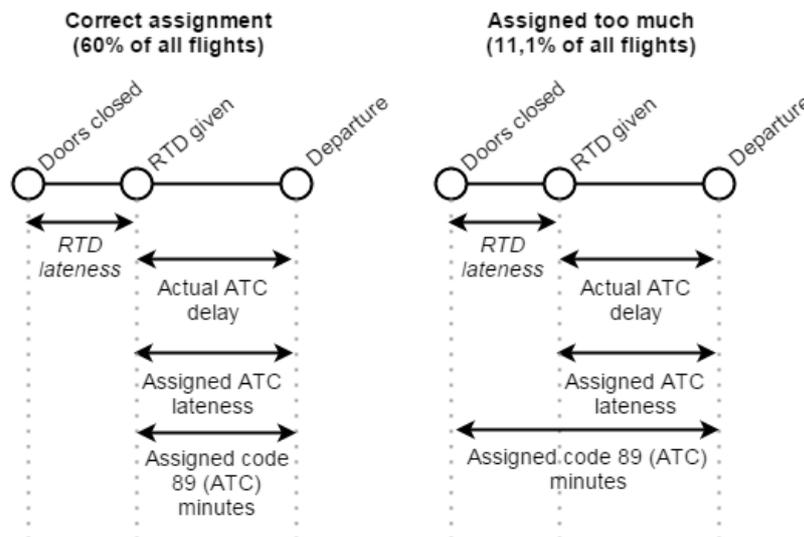
Figure 15 – Code 89 assignment, correct and incorrect

Although we found this to be true for only 400 departures, we only compared it to the moment the passenger door was closed. Of course, it is quite possible that pilots also do this with other milestones such as fuelling or pushback, further explaining the 30%. Another explanation why flights receive too few ATC delay minutes might simple be that the TLO identified a delay cause for which we do not have a measurement (as indicated in Figure 16 by a question mark), or that the departure request was given before scheduled departure. In this case, the time between these two is not accounted as ATC delay as, strictly speaking, the aircraft is not delayed. However, in this case, the aircraft could have departed earlier when ATC would have had an earlier departure slot. An argument can be made for both assigning ATC lateness for the entire time or just for the time after scheduled departure. This shows that while our ATC lateness calculation provides a good impression of local ATC performance, there are some limitations in its calculation.

However, despite these limitations, we can still draw some conclusions on these findings. First, the statement that 11% of all flights receive too much code 89 delay still holds. While pilots can request a RTD notification too early, they cannot do this retroactively. This means that in our calculation, we can attribute too many minutes to local ATC, but not too few, meaning that the value we determined can be seen as the maximum amount of delay that could be caused by local ATC. When the assigned code 89 minutes exceed this, too much delay was attributed to local ATC. At the same time, less than 1% of all flights in our dataset have a negative RTD lateness, this number would be much higher if pilots send out a RTD notification prematurely for nearly 30% of all flights. We thus expect the percentage of incorrect assignments within this 30% of all flights to
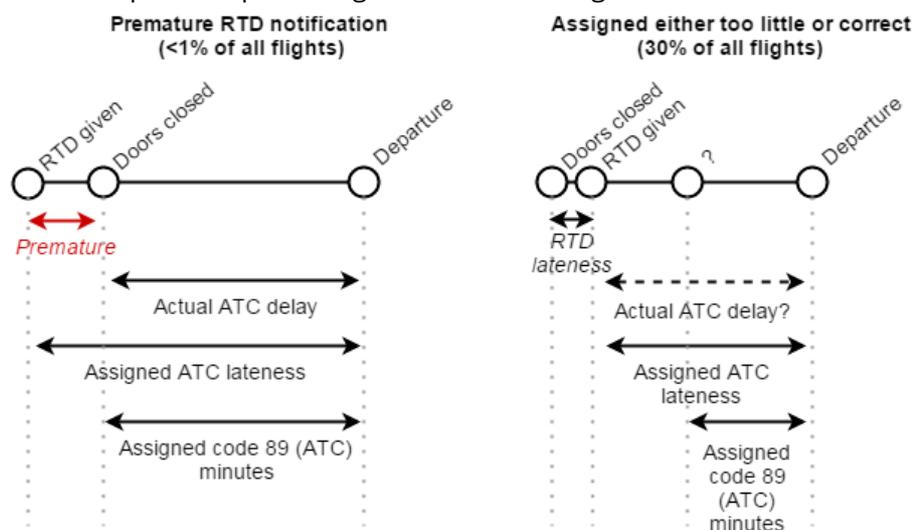
Figure 16 – Two possible explanations for too little ATC delay assignment

be significant, however, this is something we cannot quantify.

Looking at code 89 only, it is safe to say that the current delay identification method is far from perfect. For at least 10% (and presumably significantly more) of all flights, the amount of delay attributed to local ATC (code 89) is incorrect. Of course, as stated in Chapter 1, delay identification is a difficult process, as the TLO should have constant focus on every step in the turnaround to ensure proper root cause analysis, and even then, he or she still might miss vital information. Also, code 89 is the most commonly assigned code, which makes it susceptible for being a 'scapegoat' for when the TLO does not know the exact cause. Finally, currently only two codes, or causes, can be given to a delayed flight. When three or more disturbances occur, this information gets lost. All this uncertainty surrounding delay codes makes it impossible for them to be used as a reliable input in a process performance analysis.

However, we do not want to go as far as to state that the current delay identification method should be abandoned and that an automated delay code assignment would be the best option. Currently, we only have information on at what times various activities start and stop, and during which circumstances a turnaround takes place. However, there are certain causes of delay that cannot be identified from this type of data. For example, breakdown of equipment or a passenger refusing to cooperate during boarding would both not show in our data, instead making it look like a data error, resulting in a wrongfully assigned delay code if automated. An optimal delay identification process would consist of an automated delay codes suggestion that can be altered manually in exceptional cases or cases where the data simply cannot give a reliable answer.

UNIVERSITEIT TWENTE.

4.5   Conclusion

Based on the data analysis in this chapter, a few interesting conclusions can be made. First, important to note is that, while being very different techniques, the decision tree learning algorithm and delay propagation method show similar results. This means that it is safe to assume that these results show (at least for a large part) the true reasons behind departure delay.

When we start analysing the results from our decision tree learning algorithm in Table 23, we see that most of the delay originates at activities performed in the final stages of a turnaround. Scanning the boarding passes, closing the cargo doors, waiting for departure clearance from local ATC, all these activities are performed at the end of the turnaround. This makes sense, as, especially during a turnaround with a standard duration, there is some slack available between tasks. Thus, tasks at the beginning of the turnaround can be delayed without directly influencing departure punctuality. At the same time, the aircraft is scheduled to be ready for departure at D0, at which moment in time it can request departure clearance from local ATC. This means that for this event, no slack time is available and, unless previous tasks completed early, any delay in getting departure clearance will result in a delayed departure.

Next to departure-oriented tasks, one major influence on departure punctuality is arrival punctuality, as the time of arrival determines when a turnaround can start. Especially during quick turnarounds, a late arrival results in a late departure. This makes sense when considering the minimal ground time needed to service the aircraft; there is no slack to compensate for any arrival delay. During a turnaround with a standard planned ground time, any arrival delay of over 15 minutes will result in additional departure delay. This can be directly related to the minimum ground time of a quick turnaround, as the build-in slack is fully utilized by the arrival delay. To test these conclusions, we continue in Chapter 5 with a simulation study, trying to find from which improvements KLC would benefit most.

UNIVERSITEIT TWENTE.

Chapter 5: Performance Improvement Analysis

In this chapter, we test the results from our data analysis with a simulation study. By using historical data to simulate turnarounds, we create a controlled environment where we can test the impact of various changes, either in performance or in planning. To this end, we first introduce discrete event simulation in Section 5.1, after which we present our approach and model in Section 5.2. In Section 5.3, we introduce the various scenarios that we test, after which we finish this chapter in Section 5.4 with the results and implications for KLC.

5.1   Discrete event simulation

In this section we introduce discrete event simulation. We discuss the idea behind this technique, different types of implementations, and how validity and comparability between scenarios is ensured. Readers already familiar with discrete event simulation can skip to Section 5.2 where we present our implementation.

5.1.1 Discrete event simulation

A simulation model can be used for a wide variety of applications. Its main advantage is that it allows the analysis of the characteristics and performance of the system it represents, without having to change the system itself. For example, it can be used to simulate the impact of installing a new production line within a factory, without actually having to install the line. Similarly, it can determine the flight characteristics of an aircraft before it is even built. In addition, because it can recreate identical situations, it can be used to test the effect of different changes on a system and compare multiple alternatives.

A simulation model is, just like any other model, a simplified version of a certain system. Literature describes a system as "a collection of entities that act and interact towards the accomplishment of some logical end" (Winston, 2004). A system is dynamic: it status can change over time. To say something about the status of a system, we introduce the concept of a system state: "The state of a system is the collection of variables necessary to describe the status of the system at any given time" (Winston, 2004). There are two kinds of systems, discrete and continuous. In a discrete system, the state of the system changes only at discrete or countable points in time. In a continuous system, the state can change continuously. An example for a discrete system is a bank, where customers come and go, each at a different moment in time. Two customers may arrive or leave at (nearly) the same time, but no matter how close, these events are still separate. A good example of a continuous system is the heating of water. The temperature is in a constant state of change, even if a thermometer indicates a (temporary) steady state due to rounding and measurement inaccuracy. Given that a turnaround process can be recreated using discrete events, we do not discuss continuous systems any further.

A discrete event system can be simulated by two types of models: a static simulation model or a dynamic simulation model. The first represents a system at a particular point in time, the latter is the representation of a system as it evolves over time (Winston, 2004). A static simulation is often called a Monte Carlo simulation, whereas a dynamic simulation model of a discrete event system is called discrete event simulation. As we want to analyse the effect of certain changes within the turnaround process of KLC, we need a discrete event simulation model.

5.1.2 Different simulation models

Using discrete event simulation, many different models can be created. Depending on the goal and available data, each type of model has different advantages. First, we must determine whether we need a terminating or steady state model. The difference between these is whether the starting state and ending state of the process is relevant. For example, suppose we want to

UNIVERSITEIT TWENTE.   KLM

test the performance of various traffic light settings by simulating the number of cars passing a specific intersection between eight and nine in the morning. In such a model, the number of cars in cue at 7:59 and 9:01 are highly relevant for the performance, thus both starting state and ending state cannot be left out. Here, we would need a so called steady state model. This means the model starts to simulate at an earlier moment in time (e.g. at three in the morning), making sure that the state of the model at 8:00 represents a realistic value. The time preceding 8:00 is called warming up time, and is used to bring the system into its steady state. At the same time, the ending state is non-empty, there might still be cars in the cue at 9:00. As this amount of cars is relevant, this ending state should be considered when comparing performance.

On the other hand, suppose we now want test the average waiting time of customers entering a shop during an entire day. In this case, the starting state of the model is known: an empty store at opening. At the same time, the ending state is similar. At the end of the day, the shop will be empty and no more customers will be present in the shop. This is called a terminating simulation: the system is empty at the end of the simulation and the simulation stops. (Law, 2007)

Another important factor influencing our model is the type of input data. Depending on the available data, we can either use empirical (historical) data, use historical data to approximate an historical distribution, or estimate a theoretical distribution based on experience and other factors (Law, 2007). The choice between these highly depends on the data that we have to our disposal. For example, looking at our traffic light example, suppose the intersection is already in use and in the past 3 years we have gathered precise historical data of cars arriving at the intersection. In addition, we know that traffic intensity is stable and no change in traffic patterns is expected in the near future. In this case, we can use the historical data as input for our system to recreate historical situations, looking at how the system would have performed with the different traffic light settings. On the other hand, if we only have aggregated information on how many cars passed the intersection in the past few years, we can fit a statistical distribution to pick a probability of the arriving number of cars at any given moment. Depending on the number of observations and variance it can be less accurate as it might not capture complex dependencies or situations. However, the advantage over using raw historical data is that it may require less historical information and is less sensitive for data errors or outliers. The final option is using a statistical distribution not based on historical data but on estimates. This option is useful when no historical information is present or when this historical data is not representable for future situations. In the traffic light example, we might use it when the intersection is to be newly built (and no historical data is present) or a new business park just opened nearby, increasing traffic intensity. A statistical distribution is also useful when a less complex model is required or to accommodate for data points outside the historical data that might influence performance greatly. (Law, 2007)

5.1.3 Common random numbers

In Section 5.1.1, we mentioned that a simulation model can recreate identical situations repeatedly, which makes it useful for comparing the effect of different changes to a system. By having identical circumstances, we make sure that any changes in performance are caused by the different scenarios, not because of a changing situation. These identical runs are created using common random numbers, or pseudo random numbers. Unlike what the name 'common random numbers' suggests, the numbers are not entirely random, but calculated based on the previous random number that was requested. A fixed formula is used to determine the next number. This formula is formulated in such a way that there is no other relation between the numbers other than their sequence. This way, by determining a starting value (called a seed), a large list of values can be generated, which serve as input for the process. By fixing the seed value for each scenario, we ensure that identical situations are created for each scenario. Looking at our traffic light

UNIVERSITEIT TWENTE.

example in Section 5.1.2, we thus have the same number of cars arriving, each at the same moment in time. Thus, any changes in waiting time between scenarios are a direct result of the different settings of the traffic lights and not because cars were arriving at a different moment in time.

5.1.4 Number of replications required

The final important factor of simulations that we discuss is the number of replications that is required for a representative result. In Section 5.1.3, we discussed common random numbers, to take away variance between scenarios. However, between different runs within a scenario, we want to test as many input situations as possible and thus, ideally, we have to perform a run for every situation possible. However, in complex systems, this number of runs would be extremely high, and impossible to perform within a reasonable time limit. Therefore, we have to determine the number of replications that covers an adequate amount of possibilities, while at the same time does not require an extreme amount of calculation time. (Law, 2007)

Each model requires a different amount of replications, depending on the variance of both the model and the input data. There are multiple ways to estimate or calculate the number of replications required, which we do not discuss in detail. For this research, we chose the sequential procedure. The advantage of this technique is that, as long as the tested number of replications is at least 10 and the confidence level required by the researchers is more than 85%, it is a precise calculation whereas other techniques use estimations for the number of runs required. The drawback of the sequential procedure is that it requires us to perform a large number of replications to determine at what number of replications the results met the required confidence level. Because we have quite a large number of scenarios to test, performing the sequential procedure on a single scenario and a single turnaround type requires only a limited amount of calculation time compared to running an estimated 'safe' number of replications on all scenarios and turnaround types.

The sequential procedure works by looking at a single output variable. For our model, we take the departure delay (D0) of each replication. The confidence interval, average and variance of this variable are calculated incrementally after each replication. This means that for each replication, these values are calculated from the first replication until that replication. The calculated confidence interval indicates (with a certain confidence) where the 'actual average' of the variable is. Of course, we want to get a high confidence with an acceptable interval width. After each replication, the resulting confidence level is compared to the predetermined required confidence level, to see whether the required confidence level is achieved. (Law, 2007)

5.2 The simulation model

In this section, we present our simulation model. To this end, we present how we model the turnaround process into our simulations in Section 5.2.1. In Section 5.2.2, we discuss what we use as input and determine the required number of replications.

5.2.1 The model

To test the impact of the various changes we want to test, we create a simulation model of a single turnaround. In this model, we execute all tasks and events of which we have historical data. This means that we simulate a single turnaround in each replication. We do not create a model that represents an entire day, by linking multiple turnarounds and flights together. To create a realistic model of a turnaround that represents actual behaviour of parties involved, we decide to use 'activity slack' as our historical input. This activity slack can be seen as the time between the moment an activity could have started and the moment it actually started. We calculate the moment it could have started using the dependencies and planned times of events and activities.

When an activity has multiple dependencies, we take the dependency that was finished last and use this as input for our slack calculations. Three examples on how we determine slack can be found in Figure 17. The first example is a regular calculation; a task has multiple dependencies of which we take the time the last one was finished to determine the activity slack. The second example shows how we handle activities that do not have dependencies; we take the planned starting moment from the OCP-ZY20 document. The third example shows an exceptional situation, namely that activities start before their dependencies, resulting in negative slack. As discussed in Section 4.3, this might happen with the ready to depart notification. The pilots request a departure slot prematurely, which results in a 'negative' slack. As this negative slack represents actual behaviour, we decide to leave this data in our input data.
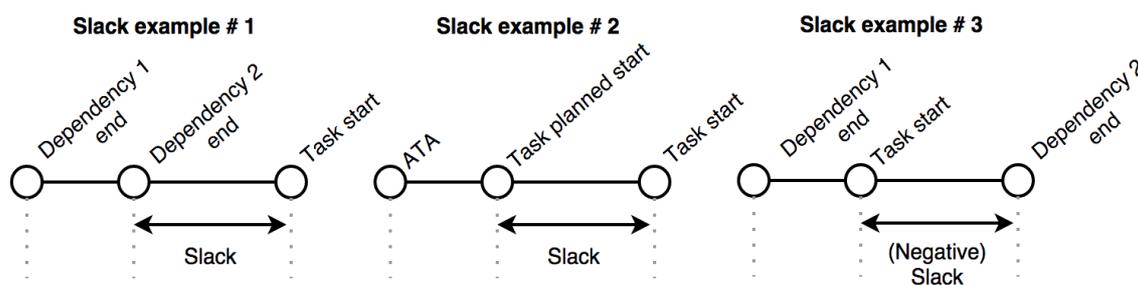
Figure 17 - Slack calculations

Each turnaround is simulated in a single replication; we do not simulate turnarounds on a schedule level over the day. This is done because we only have detailed information of turnarounds at Schiphol. All outstation turnarounds and flights would simply be an historical duration, which adds little value to our model, and might even result in less realistic results, as we do not know whether changing departure punctuality may influence flight duration for example. Therefore, we see each turnaround as a single process, starting the simulation at aircraft arrival, and ending it at departure. However, there is a limited number of operators available for each task and by recreating each simulation independently, we would lose capacity constraints between simultaneous turnarounds we could model when recreating an entire day of fleet movements. For this purpose, we only use historical data from turnarounds that took place in a similar departure bank and comparable planned ground time. This way, historical non performance caused by resource constraints is still maintained within the model.

For our simulation, we use a terminating model, with empirical data as input. As we simulate each turnaround separately, the starting state and ending state of each run are unimportant. The aircraft arrives at a certain moment in time, and departs again after all tasks have been processed. After departure, the model is empty and the simulation stops. Therefore, we do not need any warming up period.

5.2.2 Input data
As mentioned in Section 5.2.1, we use historical slack values as input for our simulation model. Because we have a sufficiently large data set of historical data, we use the specific numbers rather than approximating the historical data with an empirical distribution. Over the past 1.5 years, little has changed regarding the turnaround process that would render the data less reliable to test for future changes. In addition, no future changes are expected at this moment in time that would change the relevance of this simulation, other than that the Fokker aircraft will be replaced with Embraer E175 aircraft. This makes any results from the Fokker turnarounds less relevant, the Embraer E190 turnarounds will still be relevant in the near future.

UNIVERSITEIT TWENTE.

Just as with our data analysis, we divide our data set into smaller sets, depending on turnaround type. Activities show different behaviour between turnaround types, and also have different dependencies. For example, for a D6 gate departure, the scanning of the boarding passes is scheduled to start at D-30 minutes, whereas during a semi contact turnaround this is D-20 minutes. In addition, some activities are not present at certain turnaround types. For example, a semi contact departure does not have busses, while most Fokker platform departures take place without the need of a pushback truck.

The process of a single run of a turnaround can be found in Table 25. Most steps are self-explanatory but some might need some clarification. Step 2.1 filters for tasks that do not occur during a certain turnaround type, but also to determine whether tasks with an uncertain presence should occur or not. For example, the water replenishment task is not always executed as the water supply is regularly multi-stretched. Similarly, the Fokker can depart in a forward movement from specific locations on the platform, not needing a pushback truck. Using a probability based on historical observations a task is present during the simulated turnaround or not. If it is not to be executed, we 'schedule' it to be 'executed' right at the moment the aircraft arrives, with a duration of zero minutes. This way, any tasks that would be dependent on the not executed task can still start.

Single simulation replication step-wise approach
1. Pick a historical turnaround
1.1 Look up its planned ground time
1.2 Look up its A0
1.3 Look up its departure bank
1.4 Check whether its ground time met the minimal duration
2. For every task and event, perform the following steps:
2.1 Use the historical probability to determine whether we include it in this replication
2.2 Look up its scheduled start for this turnaround type
2.3 Pick a historical slack and duration from a comparable turnaround
2.4 Schedule it to start at its scheduled start, added with the random slack
3. Start the simulation by executing all tasks at their scheduled start
3.1 If a task cannot start due to a dependency, wait one minute and try again
4. Wait until all tasks and events are executed and calculate D0

Table 25 – Simulation steps

In step 2.3, we pick a historical value from a comparable turnaround. Here, a comparable turnaround means a turnaround of the same type, from the same departure bank (time of day) and a comparable planned ground time. This final restriction is relevant for the standard turnarounds, where planned ground time can differ greatly (see Section 2.3.2). For example, during a 3-hour turnaround the cleaning task may start 1 hour after arrival, resulting in a very large slack value. If this slack value is used for a 50-minute turnaround the cleaning task would start after scheduled departure, resulting in an extreme departure delay. A slack value taken from a 3-hour turnaround is not relevant for a 50-minute turnaround, hence the final restriction. When no comparable turnaround can be found, the bandwidth of the planned ground time is increased to allow shorter turnarounds as a source of input data until the algorithm can find an input value.

5.2.3 Number of replications
Finally, before we introduce the scenarios that we want to simulate, we have to determine the number of replications we need per scenario. As mentioned in Section 5.1.4, we use the sequential

procedure to determine the number of replications we need. To this end, we set a required confidence level of 95%, pick a single turnaround type to test (the Embraer 35 minutes departing from the Platform), and run a large number of replications (10,000). We expect 10,000 replications to be sufficiently large to cover our required confidence with some margin. This way we only need to run a single set, rather than rerunning the entire set of simulations because we were some replications short. Using the sequential procedure, we can now determine the number of replications that we need for each scenario.

Metric	10,000 replications	2,210 replications	714 replications
Average D0	9.21	9.57	9.34
Resulting confidence	97.6%	95%	92.5%
Calculation time E90 35 PL original scenario	6 hours	80 minutes	25 minutes

Table 26 – Number of replications

A summary of the results can be found in Table 26, where we see an average D0 of 9.21 minutes when using all 10,000 replications. When applying the sequential procedure on the results, we determine that using all 10,000 replications, we have a 97.6% confidence. When we reduce this confidence to 95%, we need only 2210 replications. Lowering it to 92.5% reduces the number of replications needed even further to 714. By lowering the required confidence, we reduce the required number of replications, which in turn greatly reduces calculation time. A single run of 10,000 replications of the tested scenario takes roughly 6 hours to complete. If we would only use 714 replications, we reduce the required calculation time per scenario to 25 minutes. Given the large set of scenarios we need to simulate, we chose a required confidence level of 95%, thus limiting the required number of replications to 2210 replications per scenario. This way, we still have a high confidence level of 95%, but gain a very significant reduction in calculation time. Note that calculation time might differ per turnaround type and scenario, so not all runs will take exactly 80 minutes to complete.

5.3 Scenarios

In this section, we introduce the various scenarios that we test, together with the results. To this end, we discuss the different turnaround types that we simulate in Section 5.3.1, after which we discuss the various scenarios in Section 5.3.2. The results of the simulations are presented in Section 5.3.3.

5.3.1 Turnaround types

In our input data, we have 15 turnaround types, 9 describing Embraer departures and 6 describing Fokker departures. However, we cannot simulate all 15 types using our simulation model. First, some tasks are executed multiple times during first of day turnarounds (see Section 4.1.3). Because of this, the first of day turnarounds prove to be very hard to capture in our model. The 'duration' of each task in our input data differs greatly, depending on the time the first occurrence of the task started and at what time the second occurrence ended. Getting relevant results without interfering in the process too much (limiting independency), proves to be very hard. Therefore, together with KLC, we decide to leave these turnaround types out of our scenarios. Next, a similar issue arises with the Fokker turnarounds. Because the Fokker does not support the ACARS system, we do not have any data on when the doors were closed. This means that when looking at the critical path of the turnaround process, vital information is missing. Our slack calculation depends on dependencies between tasks and when too many dependencies are missing, the slack calculation becomes unreliable. For example, during a Semi-contact turnaround, the first and only measurement before the ready to depart moment is the scan of the final

UNIVERSITEIT TWENTE.

KLM

boarding pass. Therefore, the slack for the ready to depart moment is calculated based on this value only. However, ready to depart is also dependent on other tasks, such as the closing of the cargo doors. This way, the calculated slack is less reliable for scenario testing. For example, when we want to test the influence of changing the start of the boarding pass scan process on departure punctuality, the results are only reliable when the used historical slack value was actually caused by late boarding, not when it was caused by the late closing of a cargo door. Therefore, we decide to leave the Fokker turnaround types out of our scenarios. KLC agreed with this decision, as the Fokker aircraft are scheduled to be replaced in the near future. This leaves us with six turnaround types, being all non-first of day Embraer turnaround types. These six are listed in Table 27.

Turnaround type	Abbreviation
Embraer, non-QTA, Semi contact	E90 50 SC
Embraer, non-QTA, D6-gate	E90 50 DG
Embraer, non-QTA, Platform	E90 50 PL
Embraer, QTA, Semi contact	E90 35 SC
Embraer, QTA, D6-gate	E90 35 DG
Embraer, QTA, Platform	E90 35 PL

Table 27 – turnaround types as simulation input

5.3.2 Scenarios

The final step is determining what scenarios we want to investigate. For this purpose, we have consulted KLC senior management by discussing our finding of Chapter 4 and together created a list of scenarios that we want to simulate. This list can be found in Table 28. First, the original scenario is an unaltered scenario, to provide a baseline performance, with which we can compare the performance of each scenario.

The second scenario, lengthened ground time, is a scenario where we increase the planned ground time of a QTA from 35 to 40 minutes. The idea behind it is that a planned ground time of 35 minutes might prove to be insufficient, as (in practice) there might simply too little time available to perform all tasks. By allowing five additional minutes of ground time, we would see an improved departure punctuality if this assumption holds. There are two variations, one where we do not alter the scheduling of the tasks within the turnaround process, effectively shifting all departure oriented tasks forward and creating a buffer between arrival oriented tasks and departure oriented tasks. In the second variation, we also change the scheduling of departure oriented tasks. For example, when a certain task was scheduled to start at D-10, it will now be scheduled at D-15.

The third scenario is altered arrival punctuality. By changing the arrival punctuality of the arriving flight, we test the sensitivity of the ground process on this punctuality. We create four arrival punctuality scenarios, the first one (3.1) being that every aircraft arrives at the latest at its scheduled arrival, it may arrive earlier, but not later. In scenario 3.2, aircraft always arrive 15 minutes ahead of schedule, effectively lengthening the available ground time. In scenario 3.3, we use a historical A0, but improve this by 5 minutes. Scenario 3.4 is the other way around; here we deteriorate the punctuality by 5 minutes.

In the next scenario, we test what the influence would be if we reduce the number of turnaround types within the OCP-ZY20 document. Currently, a different schedule is created for every turnaround type with slight differences between the different types. By creating just two turnaround types; one QTA and one standard, we test to see whether these scheduled times can be merged into a single schedule or that the different turnaround types actually need different schedules. To this end, we created two schedules; the first merging all schedules into the earliest

time within these schedules, the second to the last time of these schedules that is still feasible. For example, the scanning of the boarding passes starts at either D-27 at the D-gate, at D-24 for a platform departure, and D-17 for a semi-contact departure. In scenario 4.1, we created a single schedule where the scanning starts at D-27. In scenario 4.2, we set the scanning to start at D-24. Note that we do not set it to be D-17, because in that case, there would be too little time to also transport the passengers from the D6-gate to the aircraft in time.

The fifth scenario is an altered scanning process. As we found in Chapter 4, the scanning of the boarding passes is a process that often correlates with departure delay. In scenario 5.1, we therefore test the impact of starting 6 minutes earlier with the scanning of the boarding passes. In scenario 5.2, we test what would happen if we close the gate 3 minutes earlier, effectively shortening the duration of the scanning process with 3 minutes.

Important to note is that during our simulations, we do not take the arrival of the passengers at the gate into account. Because we do not have data on these arrivals, we assume that passengers are standing readily at the gate, and closing of the gate 3 minutes early should be possible without missing any passengers. In the first scenario, we presume the duration of the scanning process to be equal to a historical duration. This means that, when we start the process earlier, we can also finish earlier. Of course, in reality, the scanning process depends on the availability of the passengers: when late transfers occur, the gate cannot be closed sooner.

The final scenario that we test using our simulation model is an improvement of the local air traffic control (ATC). Using the ready to depart notification, we can see what would happen if local ATC would grant a departure slot sooner. In scenario 6.1, we test what would happen if the maximum time that an aircraft has to wait would be no more than five minutes. In scenario 6.2, we test what would happen if local ATC would grant a departure slot instantly, eliminating local ATC delay.

Scenario
1. Original
2. Lengthened ground time
2.1 QTA ground time is increased without altering OCP-ZY20 schedule
2.2 QTA ground time is increased, OCP-ZY20 optimised
3. Altered arrival punctuality
3.1 All aircraft arrive at the latest at their scheduled arrival time
3.2 All aircraft arrive exactly 15 before their scheduled arrival time
3.3 All aircraft arrive 5 minutes before their historical arrival time
3.4 All aircraft arrive 5 minutes after their historical arrival time
4. Reduced number of turnaround types to two types, QTA and non-QTA
4.1 Align all scheduled times with the earliest time of all schedules
4.2 Align all scheduled times with the latest time of all schedules
5. Altered scanning process
5.1 Start scanning 6 minutes earlier
5.2 Shorten the scanning process with 3 minutes
6. Local ATC improvement
6.1 Assign a departure slot at most 5 minutes after ready to depart
6.2 Assign a departure slot instantly after ready to depart

Table 28 – Simulation input scenarios

UNIVERSITEIT TWENTE.

KLM

5.3.3 Results

Using a 12-core CPU and creating 2200 replications per scenario, the entire model takes 12 hours and 43 minutes to complete. The results can be found in Table 29 where we listed the average departure punctuality (D0) per scenario per turnaround type. Notice that for the second scenario, the results of the standard turnarounds are missing. This is because this scenario only affects the QTA turnarounds, thus no simulation of the turnarounds with regular ground time was needed.

First, we compare our 'original' scenario with the historical data. In the original data, we did not alter any parameters within the model; it thus should give similar performance as the historical data. We note that for some turnaround types, the historical data is more closely approximated than with the others, but the differences in D0 distribution are consistent. In Figure 18, we see a comparison of the D0 distributions between the simulations and the historical data. The observed D0 is placed on the horizontal axis, the vertical axis represents the number of observations. Also, the blue line represents the historical observations, whereas the orange line represents the simulated observations. The simulated results are slightly skewed to the right, which results in a higher average departure delay. Therefore, note that the delay of the original scenario does not represent actual delay. However, as the distributions are similar, the results of the different scenarios are representable, meaning we can expect similar effects on departure punctuality in real life as we get from our simulation model.
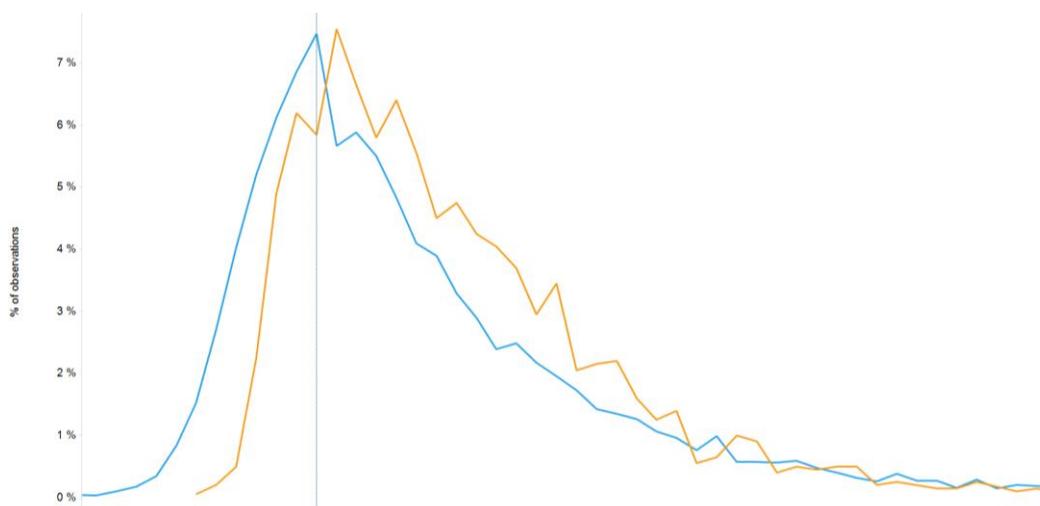
Figure 18 – D0 observations E90 35 SC, historical (Blue) versus simulated (Orange)

Lengthened ground time

First scenario to discuss is the extended ground time for the quick turnarounds. As this has no influence on the standard turnarounds, we have left them out of the table. The idea behind the extended ground time is that the current 35 minutes of planned ground time do not suffice in some circumstances and thus by extending it to 40 minutes, more flights might depart on time.

As we can see in the results, when we do not alter the scheduling of activities, average departure delay seems to decrease slightly, with the E90 35 SC type having the best average improvement with 0.47 minutes. A rather poor improvement compared to the 5 additionally available minutes. This result might be explained by looking at the scheduled times of activities. Most tasks that we identified in Chapter 4 that show correlation with departure delay are departure oriented, and take place during the final parts of the turnaround. Because their scheduled starting time is based on D-#, or '# minutes before departure', postponing the departure will actually also postpone the scheduled start of the activity. Therefore, if the delay was previously caused by the task starting too late, or taking too long, the problem will persist, it will only have shifted 5 minutes. This is

confirmed when looking at the results from scenario 2.2, where we changed the OCP-ZY20 to execute all departure oriented tasks earlier, placing the additional 5 minutes at the back of the turnaround process. Here we see a high decrease in average delay, with the E90 35 SC turnarounds having the best improvement of 3.7 minutes. The results from these two scenarios confirm that delay is mostly caused in the final stages of the turnaround process, where there is little slack to compensate any disruptions. However, the question is whether the results of scenario 2.1 would hold in the long run. If the operators found out that those additional minutes right before departure are used as slack they might simply take more time to complete their task because they know there is no direct consequence for them personally.

Scenario	E90 50 PL	E90 50 SC	E90 50 DG	E90 35 PL	E90 35 SC	E90 35 DG
1 Original	**7.45**	**6.74**	**5.09**	**9.97**	**11.51**	**7.80**
2 QTA ground time + 5 minutes						
2.1 Unaltered ZY20	-	-	-	9.83	11.04	7.53
2.2 Altered ZY20	-	-	-	6.47	7.77	5.36
3 Altered arrival punctuality						
3.1 A0 = 0	7.45	6.74	5.07	9.71	10.97	7.45
3.2 A0 = -15	7.44	6.74	5.02	9.62	10.62	7.21
3.3 A0= x + 5	7.61	6.96	5.39	10.86	12.97	8.84
3.4 A0= x – 5	7.44	6.74	5.03	9.7	10.90	7.40
4 Reduced number of turnaround types						
4.1 Earliest	6.02	4.78	5.09	8.67	10.05	7.71
4.2 Latest	7.41	5.38	6.43	9.77	10.42	8.42
5 Altered scanning process						
5.1 Start – 6	5.78	5.21	4.45	8.42	10.20	7.06
5.2 Duration – 3	5.84	5.81	4.52	8.47	10.37	7.06
6 Local ATC improvement						
6.1 ATC = max 5	6.67	6.24	4.76	9.53	11.21	7.60
6.2 ATC = 0	4.51	4.35	3.35	7.81	9.62	6.30

Table 29 – Simulation results per scenario (average delay in minutes)

Altered arrival punctuality

In order to investigate the influence of arrival punctuality (A0), we ran multiple scenarios, each with a different arrival punctuality adjustment. The first scenario, where all aircraft arrive at before or at their scheduled arrival time, shows no change in departure punctuality for the standard turnaround types. However, the quick turnaround types each show a slight improvement, again the E90 35 SC having the best improvement with 0.5 minutes. This indicates that the current minimum ground time of 35 minutes suffices in most cases, as only turnarounds with an extreme arrival delay benefit from this improvement. Similarly, an A0 of -15 minutes (thus, all aircraft arrive 15 minutes early) shows no improvement at the regular turnarounds and a slightly better improvement at the quick turnarounds. This small improvement may result from situations where an arrival oriented activity is delayed, allowing the additional 15 minutes of ground time to compensate for this delay.

A more significant influence can be seen when we deteriorate the arrival punctuality. In the "A0 + 5" scenario, we added 5 minutes of arrival delay to the historically drawn A0. This way, an aircraft previously arriving 3 minutes early, would now arrive 2 minutes late. Especially for the quick turnarounds, this seems to have a negative influence on departure delay. However, while the average delay increases, it does not increase by a full 5 minute, indicating that there was still some

UNIVERSITEIT TWENTE.   KLM

slack to be found in the process. At the same time, improving the A0 with 5 minutes seems to have a slight positive influence on D0 of the quick turnarounds, but this effect is very small.

Reduced number of turnaround types

KLC currently uses 15 turnaround types to define its turnaround processes. KLC management asked us whether they could reduce this number, without deteriorating departure delay. This means combining the various turnaround type schedules into a single schedule. Due to the fundamental differences between a standard turnaround and a quick turnaround, we decided to create a different turnaround type for each of these two, effectively reducing the number of turnaround types in our simulation model from six to two. The two different scenarios we test are described in Section 5.3.2. First, the second scenario we test shows a slight improvement for the semi contact departures, no improvement at the platform turnarounds and a deterioration of departure punctuality for the D6-gate departures. This is probably caused by the scheduled start of the boarding pass scans and the busses, in the second scenario, this is set to D-24. This is the same time as the current platform ZY20s, 7 minutes earlier than in the current semi contact ZY20s and 3 minutes later for the D6-gate ZY20s. This seems to directly influence the departure punctuality.

In the first scenario, we picked all the earliest possible times from every turnaround type, resulting in a significant improvement at both the platform and semi contact departures, but no improvement at the D6-gate departures. Again, this can mostly be explained by looking at the start of the boarding pass scanning process. In this scenario, we start a full 10 minutes earlier during semi contact turnarounds, and 3 minutes earlier during platform turnarounds. The D6 gate showed little to no improvement, as the created turnaround schedule is nearly identical to the current D6-gate schedule.

Altered scanning process

The fifth scenario is an altered scanning process, where we test whether an earlier start or shorter duration of the scanning process improves departure punctuality. Given that the boarding pass scanning process is an important variable in the results of our data analysis, we expect these changes to have a positive impact on departure punctuality. The results of our simulation confirm this expectation. Especially at the platform and semi contact departures, the improvements are significant; both for starting earlier and stopping earlier. The D6-gate departures already start very early in the process at D-30, starting even earlier seems to have less effect than it has on the other turnaround types. Interesting to see, is that the semi contact departures do not benefit from a shorter duration of the scanning process as much as the platform turnarounds do. This might be explained by looking at the scanning duration distribution of the two departure types, which can be found in Figure 19. Here, we see that the scanning process at the semi-contact departures (orange) is clearly shorter than at the platform departures (blue). This explains why the semi-contact departures are less sensitive to a shorter duration of the scanning process, because (on average) it is already shorter, making it even shorter is of no influence. A logical explanation to why the semi-contact departures have a shorter average duration can be found in the time window available for the scanning of the boarding passes. At the semi-contact departures, the process is scheduled to start at D-17, whereas for platform turnarounds, this is D-24. However, the
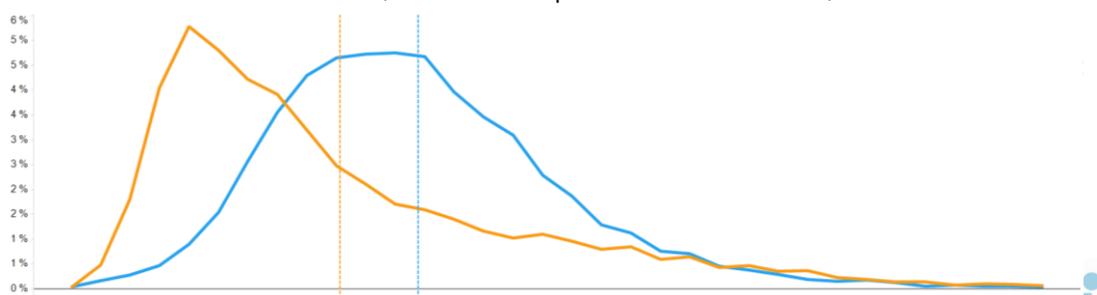

Figure 19 – Duration of the scanning process, Semi-contact (orange) versus platform (blue)

respective planned completion time is D-8 and D-9. This means that for semi-contact departures, the available time window is shorter, thus it is no surprise that the actual average duration is also smaller.

The results of this scenario are in line with the results from the scenario where we tested a reduced number of turnaround types. There, we saw similar results in the scenario when during platform and semi contact departures the scanning process started earlier.

Local ATC improvement

The final scenario we test is a potential improvement of Local ATC performance. First, we test what would happen if the local ATC delay was at most 5 minutes. Next, we test for what would happen if there were no delay at all. Note that we define ATC delay as the time between the completion of all activities and the moment of actual departure. These activities include fuel, pushback, doors and the ready to depart notification. Of course, it might be possible that any delay occurring between these moments and actual departure occasionally has a cause other than Local ATC, such as the sudden breakdown of the pushback truck. Therefore, while it is safe to assume most 'ATC delay' is actually caused by Local ATC, we must emphasise that more information is needed on these occasional causes of delay to make a fair judgement.

That being said, the results of the ATC improvement scenario show a strong improvement of departure punctuality for all turnaround types. In scenario 6.1, where we limit the amount of delay to a maximum of 5 minutes, the improvement is not as high, with an average improvement of 0.5 minutes. In scenario 6.2 however, where we eliminate ATC delay entirely; we see an average improvement of 2.1 minutes. Thus, the simulation results confirm our findings in Chapter 4, where we also found local ATC to be of major impact. This actually has a logical explanation. The current turnaround process is designed to be completed at D0, or scheduled time of departure (STD). All activities and events are thus focussed on this moment in time. When an operator has some spare time within his time window, he might take it easy, knowing that he still has some time left, and the aircraft will not depart late when he takes the time he was originally assigned. Therefore, in general, the ground process finishes before STD only by exception, absorbing any slack that was present in the ground time. The process is designed for local ATC to give a departure slot instantly; there is no build-in slack to compensate for local ATC delay. Therefore, if local ATC cannot give such a slot instantly, it will result in a delayed departure.

5.4   Conclusion

In general, the results of our simulations confirm the results of the data analysis of Chapter 4. Using our simulation model, we were able to test six scenarios to verify the impact of certain changes, suggested by both the results of Chapter 4 and KLC Senior management. First conclusion is that increasing the planned ground time for QTAs from 35 minutes to 40 minutes seems to result in a serious improvement of departure punctuality. However, this only has an effect when the departure oriented tasks are rescheduled to take place earlier in the process, thus creating a 5 minute buffer at the end of the turnaround process. If the departure oriented tasks are scheduled at their original time (thus creating a 5 minute buffer in the middle of the turnaround process) had very little effect. Therefore, adding 5 minutes of ground time is only useful when placed at the end as a buffer. This is in line with the findings of Chapter 4, where we conclude that most delay is mostly caused by disruptions in the final stages of the turnaround process. However, adding such a buffer might be less effective in the long run. Both operators and their operational schedulers will find out that these additional 5 minutes of ground time are added as a buffer and might claim them prematurely by stretching the time they need for their task or solve capacity constraints by rescheduling tasks more closely to scheduled departure, nullifying the effect of the added ground time.

UNIVERSITEIT TWENTE.   KLM

However, besides adding 5 minutes of ground time, other changes in the current turnaround schedules show a more promising improvement. The results of scenarios 4.1, 5.1 and 5.2 show that by starting the boarding card scanning process earlier, a significant departure punctuality improvement can be expected, especially at flights departing from the platform or semi contact. On a schedule level, arrival punctuality should be maintained at the current level. Improving it by 5 minutes showed little improvement, whereas deteriorating it with 5 minutes resulted in significant deterioration of the departure punctuality. Finally, the biggest improvement can be gained by minimizing local ATC delay. The simulations show that by minimising the time between 'ready to depart' and actual departure, a significant gain in departure punctuality could be achieved. In Chapter 6, we discuss what these results imply for KLC and what improvements will increase departure punctuality the most.

UNIVERSITEIT TWENTE.

KLM

Chapter 6: Conclusions & Recommendations

In this final chapter, we summarize the results from the previous chapters to find an answer to our main research question. To this end, we first do a short recap of our work and answer the sub research question in Section 6.1, after which we answer our main research question and discuss the implications for KLC in Section 6.2. In Section 6.3, we discuss the limitations of this research, and do some suggestions for further research.

6.1   Answering of the sub research questions

This section is meant as a summary of the results of Chapters 2, 3, 4 and 5 and the answers to the research questions they provide. First, we look back at the current performance measures of Chapter 2 (Section 6.1.1), after which we discuss our data gathering process of Chapter 3 (Section 6.1.2). In Section 6.1.3, we discuss the data analysis of Chapter 4. Finally, in Section 6.1.4, we discuss our simulation and its results.

6.1.1 What is the required performance level of the turnaround process of KLC at Schiphol airport?

KLC uses many key performance indicators (KPIs) to get insight into its general performance. Each day, the overall results of the previous day are discussed, using delay codes to identify delay sources. These delay codes are also used to discuss performance during monthly performance meetings with the various service providers, comparing their actual performance (measured in number of assigned delay codes) to the agreed service levels in the service agreements. KLC uses procedures for 15 different combinations of turnaround circumstances, for each turnaround type defining at what time each activity or event should take place. These 15 turnaround types are based on aircraft type, planned ground time, aircraft location and gate location.

Thus, KLC has a thorough overview of how the turnaround process should happen, and has made agreements (both qualitative and quantitative) with its service providers towards this goal. KLC currently uses delay codes to identify causes of non-performance and gain insight into service provider performance. However, while this may give a general impression of performance, it is by no means an extensive tool to monitor the actual performance. Non-performance is only registered in a delay code when it has caused the aircraft to be delayed and, more importantly, when the TLO has identified the correct cause.

6.1.2 What data is required to assess the actual performance of the turnaround process?

For this research, we gathered a large set of data from the AP2K project, describing various aspects of the turnaround process of KLC at Schiphol Airport over the course of 1.5 years. This data was supplemented with additional sources, such as weather information, and communication between the cockpit and the tower. Next, we filtered this data to include only turnarounds with reliable data. This resulted in 47,000 turnarounds out of the total set of 70,000 to be kept; the rest (roughly one third of the data) was discarded. As most of the activity related data is generated by hand, many errors (intentionally and unintentionally) made their way into our input data. Most turnarounds were discarded, because these activity related measurements were missing or unreliable. Deleting a third of our data means that the relevance and representability of our research is lowered. However, a small analysis (Section 4.3.3) showed that there are only minor differences in the results of the delay propagation method when comparing the filtered data set and the unfiltered data set. This suggests that the characteristics of the filtered and unfiltered data are very similar, and thus the results of this research are both representative for both the filtered data and the unfiltered data. However, to increase credibility and acceptance by the business, data quality should be improved for future research.

UNIVERSITEIT TWENTE.

6.1.3 What is the actual performance of the turnaround process?

In Chapter 4, we performed our data analysis. To this end, we used decision tree learning, which is a machine learning technique. Using the filtered data from Chapter 3, we gained a set of variables that correlate with departure delay; these variables can be found in Table 30. At the same time, the delay propagation method was used to gain an alternative insight into departure delay causes. The delay propagation method provided similar and complementary results when compared to the decision tree learning results. This means that it is likely that these variables are indeed the actual causes of delay.

Analysing these results, we see that most of the activities involved in departure delay are departure oriented and take place in the final stages of a turnaround. This makes sense when looking at the scheduled times and available slack, as there is little to no slack at the end of the turnaround. Where activities taking place during the early stages of a turnaround still have some slack between the arrival-oriented and departure-oriented activities to compensate for small disturbances, the activities at the end of the process do not have this luxury. The final passenger should be aboard at 5 minutes before departure in order to get him to his seat, stow his luggage, and fasten his seatbelt before the aircraft can depart. However, the closing of the passenger door is also scheduled at D-5, effectively creating no slack at all. This is true for all departure-oriented activities and means that, unless previous tasks complete early, any delay during the activity results in a delayed departure.

Variable	Description
Local ATC lateness	Time between aircraft ready and actual departure
Passenger door closed and cargo door closed lateness	Time at which the various doors were closed
Ready to depart lateness	Time between doors closed and ready to depart notification
Arrival punctuality	Time the inbound flight was delayed
PaxScan Dminus	Time between last boarding pass scan and scheduled departure
Pushback start Dminus	Time between the start of the pushback task and scheduled departure
Load factor	Percentage of seats occupied

Table 30 – Important activity related variables and circumstances for departure delay

Besides departure-oriented activities, one major influence on departure punctuality is arrival punctuality. Especially during quick turnarounds where available ground time is reduced to a bare minimum and thus the ground process cannot make up for the delay, a late arrival is a guarantee for a late departure.

6.1.4 Which improvements would contribute to a better D0 performance the most?

Looking at the answer to the previous question, we notice that most activities are under indirect operational influence of KLC. Arrival punctuality is primarily determined by departure punctuality at the airport the flight originates, which is primarily determined by the performance of the local ground services party. Similarly, local ATC lateness is caused by congestion. Because the large majority of flights departing from Schiphol are KLM, congestion is the result of a non-optimal network schedule. Especially during peak hours, a departure slot may not be instantly available and the aircraft might have to wait at the gate. The load factor might be under direct control by KLC, by limiting the number of seats available for booking, but deliberately flying with a half-empty aircraft is not profitable, thus not an option.

UNIVERSITEIT TWENTE.

To see which adjustments show the highest potential improvement of departure punctuality, we performed a simulation study of the turnaround process where we test various scenarios of improvement. The different scenarios and their results are summarized in Table 31. The biggest gains can be achieved when introducing some buffer time between the scheduled completion of the tasks and scheduled departure. Depending on the turnaround types, average departure punctuality increased by over 3.5 minutes. Also, by minimizing local ATC lateness the average departure punctuality can, depending on the turnaround type, be increased by up to 2.5 minutes. Next, the altered boarding pass scanning process shows high potential, allowing platform turnaround to depart 1.5 minutes earlier on average.

	Scenario	Results with regard to departure punctuality
1	Original	-
2	Increased ground time for QTAs	Only a slight improvement when used as buffer between arrival and departure oriented tasks, major improvement if used as a buffer right before STD.
3	Altered arrival punctuality	Slight improvement when arriving 5 minutes earlier, Serious deterioration when arriving 5 minutes later.
4	Reduce 15 turnaround types to 3, QTA, standard E90 and standard F70	Depending on chosen new schedule, serious improvement at platform and semi contact turnarounds.
5	Altered boarding pass scanning process	Improvements at all turnaround types for both earlier start and shorter duration. Biggest gain at platform and semi contact turnarounds.
6	Local ATC improvement	Serious improvement at all turnaround types when reducing lateness to zero minutes.

Table 31 – Summary of simulation results

Reducing the number of turnaround types shows, dependent on the chosen new schedule, an improved departure punctuality at the platform and semi contact turnarounds. The performance of a single combined schedule actually comes down to a single choice: when to start with the boarding pass scanning process. During D6-gate departures, platform departures and semi contact departures, this currently is at D-27, D-24, and D-17 respectively. When setting the combined schedule to D-27, the platform and semi contact departures will have an increased performance, when setting it to D-24, the D6-gate departures will have a deteriorated performance. Setting it to D-17 results in a very high increase in departure delay at the D6-gate and platform departures and is no viable option.

Interestingly, only a small improvement can be expected when raising arrival punctuality. This might be explained by looking at the arrival punctuality over the past 1.5 years. Roughly 75% of all flights arrived on time, with an average A0 of -4 minutes. This means that most aircraft already arrive on time at Schiphol Airport. Improving arrival punctuality thus only has an influence on a small portion of the flights and any changes thus have low average impact. For this reason, the scenario where all aircraft arrive 5 minutes earlier performs slightly better, but still no significant improvements are achieved. More interestingly, we see that deteriorating arrival punctuality shows a significant reduction in departure punctuality for some turnaround types. This all indicates that while arriving late will cause departure delay, arriving early will not make you depart early as well.

The final scenario to discuss is the increased ground time for quick turnarounds. Depending on the scheduled start of the departure oriented tasks, this can have a very high impact on departure punctuality. We already found out that most activities that cause delay take place at the final

UNIVERSITEIT TWENTE.

KLM

stages of the turnaround. These activities are scheduled at a certain number of minutes before departure (e.g. D-25). When we shift the scheduled departure forward in time, we also shift these activity forwards, effectively creating slack time in the middle of the turnaround. This shows little to no improvement in departure punctuality. However, if we adapt the schedule slightly to start departure oriented tasks earlier before departure, we see a high increase in departure punctuality. By changing their scheduled starting time, we create slack right before departure, allowing for all disturbances (including ATC lateness) to be compensated. Therefore, increasing the ground time to 40 minutes is only useful when also adapting the activity schedule.

6.2   Main research question and recommendations

Now that we have answered all of our sub research questions, we use this section to answer our main research question. We discuss the potential improvements that resulted from our simulation results, and present some recommendations on where we think KLC should focus its efforts if it wants to improve departure punctuality.

Where can KLM Cityhopper best focus its efforts to improve departure punctuality?

In Section 6.1.4, we presented a list of potential improvements, together with their expected gains. However, we did not take into account what efforts KLC would need to make in order to implement these changes. Our simulations show that to improve its departure punctuality, KLCs best option is adding a five minute buffer at the end of each turnaround. However, we understand that this is easier said than done. First of all, the operators and operator schedulers should respect this five minutes and not use it prematurely for their own interest. To prevent this, KLC can make new agreements with its service providers and monitor operator behaviour by using the newly implemented AP2K dashboard. Also, as the entire boarding process should start five minutes earlier, passengers should also be present five minutes earlier. This can be solved by providing passengers with gate opening and gate closing information, rather than scheduled departure. Finally, the additional five minutes could cause problems with regard to aircraft utilisation on a network level. Because turnarounds require more time, there might be insufficient aircraft fulfil all flights during the day. However, if these three problems can be addressed, departure punctuality will increase greatly when a ground time buffer is introduced.

Next, KLC can try to reduce local ATC lateness. We suggest two options: the most straightforward option is to address the LVNL and try to find a solution together. This might include an analysis into in what circumstances trigger local ATC lateness most frequently or if there is 'low hanging fruit' that can be implemented to reduce local ATC lateness on short term. A project called CDM (Collaborative Decision Making) is currently being implemented and aims to improve communication between the tower and the cockpit, for example by reporting when the pilots expect to be ready. The second option to reduce local ATC lateness is changing the current network schedule. As KLM and KLC are responsible for the vast majority of flights departing from Schiphol, congestion is primarily caused by themselves. If flights would be more spread out over time, congestion would be less of an issue. However, this will result in lengthened transfer times, lengthened working hours for crews, and (for evening departures) risk of a closed outstation when the aircraft is delayed. However, if the resulting schedule is more robust, it might be worthwhile to implement.

Another option to improve departure punctuality is to investigate the possibility of starting the boarding pass scanning process earlier. Especially at platform departures, we see significant improvements if the scanning process starts at D-30, rather than D-25. Of course, such an adjustment is not without complications. Availability of the gate agents and busses must allow it,

and passengers have to show up at the gate earlier as well, which might be impossible if they have a tight transfer. However, even if some passengers arrive late, the effect might still be positive. As most people are scanned earlier, the first bus with passengers can depart sooner. This results in that up to 50 passengers (the first bus) will be able to board earlier, enlarging the time window available for the passengers to find their seat and stow their luggage and improving the 'cabin ready' moment.

Next, we recommend redefining the current OCP-ZY20 turnaround types. We investigated the projected impact on departure punctuality if we were to reduce the number of turnaround types to three, a quick turnaround, a standard Embraer E190 turnaround and a standard Fokker F70 turnaround. If the above recommendation with regard to the boarding pass scanning process can be implemented, we recommend merging the turnaround types and simply use the current D6-gate schedule for all turnaround types. This improves departure punctuality at the semi contact and platform departures while at the same time reducing complexity.

6.3　Limitations & further research

In this section, we discuss the limitations of our research and present some topics for additional research with which we think KLC can gain additional improvements.

Limitations

The most important limitation of our research is the quality and availability of performance measurements. A large portion of our data is generated by hand by the operators executing the various tasks. This is a serious risk for the reliability of the results, as this way operators can (intentionally and unintentionally) influence the observed performance. For example, an operator may have forgotten to push the button when he started his activity. He discovers this when he wants to finish his task, and pushes the button twice to start and immediate end it. This is an obvious data error, which we can filter easily. However, when he discovered it 4 minutes into his task and pushed the button at that moment in time, it becomes very hard for us to track this error. In addition, manually generated data can also be influenced intentionally, for example by ending the task earlier. This way, an operator seems to have performed his task quicker. He then can blame a relative late start at his next task to traffic congestion driving there. This type of data error is hard to notice at individual turnaround level and will probably also be present in our other manually generated data.

Next, we focussed entirely on the turnaround process at Schiphol Airport; the performance of the various activities might be very different on the outstations. In addition, there are multiple service providers based on Schiphol Airport and other airlines using KLM ground services as their service provider at Schiphol Airport may have contracted other service agreements. Therefore, the results of this research are only representative for KLM Cityhopper at Schiphol Airport, and no other combination of airline and airport.

Finally, our simulation model depends heavily on the various dependencies that exist between tasks. We have gathered these dependencies to the best of our abilities by informing with experts within KLC and monitoring the process. However, although unlikely, it might be possible that there are hidden dependencies between activities or circumstances that we did not model into our simulation. These hidden dependencies may not show in the results of certain scenarios, whereas they might show when implementing the changes in real life. Therefore, any changes should be tested on a small scale first, before implementing them KLC-wide.

Further research and recommendations

UNIVERSITEIT TWENTE.

Finally, we present some topics that KLC should perform further research on, to gain more control over its departure punctuality. The first and most important topic is the number of busses used to transport passengers from the gate to the aircraft. Currently, KLC uses two busses from both the B-pier as well as the D6-gate, unless more than 70 passengers are expected at the D6-gate, in which case three busses are used. This third bus allows more spread in the arrival of the passengers at the aircraft, enabling a smoother boarding process. We performed a small initial data analysis into the departure punctuality for turnarounds departing from the D6-gate. We compared the performance of departures where two busses were used with departures where three busses were used. This showed that departures where three busses were used had an average D0 that was 1.5 minutes lower when compared with turnarounds with two busses. Further research into confirming this find is recommended and when successful, a pilot to use three busses at the B-pier might be worthwhile.

Next, in this research we use the 'ready to depart' notification to determine local ATC lateness. However, more data regarding the communication between the cockpit and the tower should be available. By examining this data, KLC might gain valuable insights into how it can reduce delay caused by waiting for departure clearance.

The final recommendation is the usage of additional measurements. As mentioned numerous times in this report, manually generated data is unreliable. However, using the ACARS system, much more information can be gathered from the aircraft itself, for example the opening and closing of fuel hatches. These measurements allow us to see the moment the fuel operator opened and closed them, meaning a more precise measurement of the fuelling process. Precise loggings of the ACARS system are unknown to us, but we expect much information to be available. Also, additional measurements regarding the boarding process would greatly increase insights into the departure process. Currently, the only information we have regarding passengers is the scanning of the boarding passes and the closing of the passenger door. If information regarding the boarding process would be available (e.g. gate agent present at gate, first passenger aboard, last passenger aboard, all luggage stowed away, all passengers seated and ready) we gain valuable additional insights into possible causes of departure delay.

UNIVERSITEIT TWENTE.

Bibliography

AhmadBeygi, S., Cohn, A., Guan, Y., & Belobaba, P. (2008). Analysis of the potential for delay propagation in passenger airline networks. *Journal of Air Transport Management* , 221-236.

Allison, P. (2009). Missing Data. In R. E. Millsap, *SAGE Handbook of Quantitative Methods in Psychology* (pp. 72-89). Thousand Oaks, CA, USA: Sage Publications Inc.

Cook, A., Tanner, G., & Lawes, A. (2012). The Hidden Cost of Airline Unpunctuality. *Journal of Transport Economics and Policy* , 157–173.

Flynn, B., Sakakibara, S., & Schroeder, R. (1990). Empirical research methods in operations management. *Journal of Operations Management* , 250-284.

Graham, J. (2009). Missing Data Analysis: Making It Work in the Real World. *Annual Review of Psychology* , 549-576.

Hsiao, C.-Y., & Hansen, M. (2006). *An Econometric Analysis of US Airline Flight Delays with Time-of-Day Effects.* Berkeley: National Center of Excellence in Aviation Operations Research.

Kohla, N., Larsen, A., Larsen, J., Ross, A., & Tiourine, S. (2007). Airline disruption management - Perspectives, experiences and outlook. *Journal of Air Transport Management* , 149–162.

Law, A. M. (2007). *Simulation Modeling & Analysis.* Tucson, Arizona, USA: McGraw-Hill.

Mladenic, D. (2005). Feature Selection for Dimensionality Reduction. *Subspace, Latent Structure and Feature Selection* (pp. 93-102). Bohinj, Slovenia: Springer.

Oshiro, T., Perez, P., & Baranauskas, J. (2012). How Many Trees in a Random Forest? *International Conference on Machine Learning and Data Mining* (pp. 154-168). Springer.

Reif, D., Motsinger, A., McKinney, B., Crowe Jr, J., & Moore, J. (2006). *Feature Selection using a Random Forests.* IEEE.

Tu, Y., & Ball, M. (2008). Estimating Flight Departure Delay Distributions - A Statistical Approach With Long-term Trend and Short-term Pattern. *Journal of the American Statistical Association* , 112-125.

van der Aalst, W. (2011). *Process Mining.* Heidelberg: Springer.

Winston, W. L. (2004). *Operations Research.* Belmont California, USA: Thomson Brooks/Cole.

Witten, I., & Frank, E. (2000). *Data Mining.* San Diego: Morgan Kaufmann Publishers.

Wu, C., & Craves, R. (2004). Modelling and Simulation of Aircraft Turnaround Operations. *Transportation Planning and Technology* , pp. 25–46.

Wu, C.-L. (2008). Monitoring Aircraft Turnaround Operations - Framework Development, Application and Implications for Airline Operations. *Transportation Planning and Technology* , 215 - 228.

Wu, C.-L., & Caves, R. E. (2003). Flight Schedule Punctuality Control and Management: A Stochastic Approach. *Transportation Planning and Technology* , 313–330.

UNIVERSITEIT TWENTE.

Wu, C.-L., & Caves, R. E. (2003b). The punctuality performance of aircraft rotations in. *Transportation Planning and Technology* , 417–436.

Wu, C.-L., & Truong, T. (2014). Improving the IATA delay data coding system for enhanced data analytics. *Journal of Air Transport Management* , 78-85.

UNIVERSITEIT TWENTE.

KLM

Appendices

Appendix A – Glossary

TLO	Dutch: *Team Leider Omdraai*. Person managing the turnaround, present and responsible the entire on-blocks time.
Hub-and-Spoke network	Plan of flights where all destinations are served from a single airport
Flight Operations	Department of KLC responsible for all flight operations
KLC	KLM Cityhopper
KLM	Royal Dutch Airlines (Dutch: "Koninklijke Luchtvaart Maatschappij")
Loadsheet	List of all people and items on board of the aircraft.
Multi-stretching	Skipping an activity between flights, e.g. no water replenishment
OCP-ZY20	KLC Procedure describing the turnaround process
Outstation	Airport to which KLC maintains a flight connection from Amsterdam
Terminal weather	The local weather at the airport
Platform	Section of the airport where aircraft are parked during a turnaround if not at a gate.
Registration change	The assignment of a different aircraft to perform a certain flight.
STA / ATA	Scheduled time of arrival / actual time of arrival
STD / ATD	Scheduled time of departure / actual time of departure
PDO / PDC	Passenger door open / passenger door closed
CDO / CDC	Cargo door open / cargo door closed

UNIVERSITEIT TWENTE.

Appendix B

OCP-ZY20 - List of activities during a turnaround at AMS of an Embraer 190, taking 35 minutes and departing from the platform

Removed due to confidentiality

UNIVERSITEIT TWENTE.

Appendix C
List of delay codes used by KLC

Number	Explanation	Department
1	ATC Delay caused by Airline Activity	ATC Local ZV
2	System Generated Code	Initial code cannot be traced ZZ
3	Missing Checked-in Passenger	PH KLC ZY missing checked in pax (ZE)
4	Priority Clashes Between KLM Deptmts.	Suppliers general ZY
5	Pre-announced Short Turnaround failed	A/c t/a services-Ramp handling ZY
6	No Gate Avail due own Airline Activ.	Airport Facilities ZY
7	Reserved -07-	
8	Startup Delay due Adverse Weather	ATC Local ZV
9	Sched Ground Time less than Decl Min	Connections & Rotations ZU
10	Excessive Hand Baggage at Gate	PH KLC ZV (ZE)
11	Late Check-in Acceptance after deadl.	PH KLM (Check-in) ZY (ZE)
12	Late Check-in Congestion	PH KLC ZV (ZE)
13	Check-in Errors, Passengers & Baggage	PH KLM (Check-in) ZY (ZE)
14	Oversales, Booking Errors	PH KLC ZV (ZE)
15	Boarding, Discrepancies & Paging	PH KLC ZV (ZE)
16	Commercial Publicity/Pax Convenience	PH KLC ZV (ZE)
17	Catering Order Late/Incorrect	Catering ZSY
18	Baggage Processing in Sorting Area	Baggage services ZY
19	Passenger with reduced mobility (prm)	PH KLC ZV (ZE)
20	BASS Disturbances	Baggage services ZY
21	Cargo Documentation/Packing	Cargo/Mail ZY
22	Cargo Late Positioning	Cargo/Mail ZY
23	Cargo Late Acceptance	Cargo/Mail ZY
24	Cargo Inadequate Packing	Cargo/Mail ZY
25	Cargo Oversales/Booking Errors	Cargo/Mail ZY
26	Cargo Late Preparation	Cargo/Mail ZY
27	Mail Documentation/Packing	Cargo/Mail ZY
28	Mail Late Positioning	Cargo/Mail ZY
29	Mail Late Acceptance	Cargo/Mail ZY
30	Loading/Unloading not Controllable	A/c t/a services-Ramp handling ZY
31	Aircraft Documentation Late/Inaccur.	Load planning department ZY
32	Loading/Unloading, Bulky, Special etc	A/c t/a services-Ramp handling ZY
33	Loading Equipment Lack or Breakdown	A/c t/a services-Ramp handling ZY
34	Servicing Equipment Lack or Breakdown	A/c t/a services-Ramp handling ZY
35	Aircraft Cleaning	Aircraft Cleaning ZY
36	Fuelling/Defuelling	Fuelling Department ZY
37	Catering Delivery Late/Incorrect	Catering ZSY
38	Lack of ULD Equipment	Cargo/Mail ZY
39	Towing/Pushback Staff or Equipment	Push-back / Support equipment ZY
40	Reserved -40-	
41	Aircraft Defects or Urgent Repairs	Technical ZT
42	Scheduled Maintenance Late Release	Technical ZT
43	Non-scheduled Maintenance	Technical ZT
44	Spares and Maintenance Equipment	Technical ZT
45	A.O.G. Awaiting Spares	Technical ZT
46	Aircraft Change for Tech Reasons	Technical ZT
47	Stand-by Aircraft, lack of planned	Technical ZT
48	Scheduled Cabin Adjustments	Technical ZT
49	Reserved -49-	
50	Reserved -50-	

Number	Explanation	Department
51	Damage During Flight Operation	ATC en route ZV
52	Aircraft Damage during Ground Oper	A/c t/a services-Ramp damage ZY
53	Reserved -53-	
54	Reserved -54-	
55	Departure Control System	Suppliers general ZY
56	Cargo Preparation/Documentation	Cargo/Mail ZY
57	Flight Plan N/A or Received Partially	Flight plan/fuel ZV
58	Other Automated Systems	Other Automated systems ZY
59	Reserved -59-	
60	Reserved -60-	
61	Flight Plan Late Completion or Change	Flight plan/fuel ZV
62	Operational Requirements, Extra Fuel	Flight plan/fuel ZV
63	Late Crew Boarding or Dep Procedures	Cockpit Crew ZV
64	Flight Deck Crew Shortage	Cockpit Crew ZV
65	Flight Deck Crew Special Request	Cockpit Crew ZV
66	Late Cabin Crew Boarding or Dep proc.	Cabin Crew ZS
67	Cabin Crew Shortage	Cabin Crew ZS
68	Cabin Crew Special Request	Cabin Crew ZS
69	Captain Request for Security Check	Cockpit Crew ZV
70	Reserved -70-	
71	Weather at Departure Station	Weather ZZ
72	Weather at Destination Station	Weather ZZ
73	Weather en-route or Alternate Station	Weather ZZ
74	Reserved -74-	
75	De-icing/Snow Removal from Aircraft	De-icing ZY
76	Snow/Ice/Water/Sand Removal from Airp	Airport Facilities ZY
77	Groundhandling imp by Adverse Weather	Aircraft loading hampered by weather
78	Reserved -78-	
79	Reserved -79-	
80	Reserved -80-	
81	ATFM due ATC en-route Demand/Capacity	ATC en route ZV
82	ATFM due ATC Staff/Equipment en-route	ATC en route ZV
83	ATFM due to Restr. at Dest. Airport	Airport Facilities ZY
84	ATFM due to Weather at Destination	Airport Facilities ZY
85	Mandatory Security	Airport Facilities ZY
86	Immigration/Customs/Health	Airport Facilities ZY
87	Airport Facilities	Airport Facilities ZY
88	Restriction at Airport of Destination	Airport Facilities ZY
89	Restriction at Departure Airport	ATC Local ZV
90	Reserved -90-	Airport Facilities ZY
91	Load Connection	Connections & Rotations ZU
92	Through Check-in Error	Through check-in error ZY (ZE)
93	Aircraft Rotation (Late Incoming AC)	Aircraft Rotation ZZ
94	Cabin Crew Rotation	Connections & Rotations ZU
95	Flight Deck Crew Rotation	Connections & Rotations ZU
96	Operations Control	Re-routing/diversion of the flight ZZ
97	Industrial Action Within Own Airline	Industrial Actions ZZ
98	Industrial Action Outside Own Airline	Industrial Actions ZZ
99	Not elsewhere specified	Industrial Actions ZZ

UNIVERSITEIT TWENTE.

Appendix D

The weekly report regarding ground process performance as created by the Manager Operational Processes & Planning to be used by the Director Ground Services & Cargo Operations KLC. (2 pages)

Removed due to confidentiality

Appendix E

Monthly non-performance overview Ground Services (ZY) on all stations, also created by the Manager Operational Processes & Planning to be used by the Director Ground Services & Cargo Operations KLC.

Removed due to confidentiality

Appendix F

Monthly non-performance comparison of Ground Services at Schiphol, non-performance compared to same month last year. Again created by the Manager Operational Processes & Planning to be used by the Director Ground Services & Cargo Operations KLC.

Removed due to confidentiality

Appendix G

The weekly report regarding ground process performance in Amsterdam, supplemented with specific service provider activity overviews, created by the various Duty Managers Ground Services. Below the specific performance chart for the water and toilet service for week 3 in 2015.

Removed due to confidentiality

UNIVERSITEIT TWENTE.

Appendix H

Monthly summary of task conformance during turnarounds, as monitored by the Duty Managers Ground Services. Summary is in shape of Excel file, with performance divided over tabs, each area of attention has a separate tab. Below are the general tab and loading-specific tab for April 2015.

Removed due to confidentiality

Tab with Loading-specific performance:

Removed due to confidentiality

UNIVERSITEIT TWENTE.

Appendix I
Performance overview KLC Flight Operations (ZV) used by (among others) Director Flight Operations.

Removed due to confidentiality

Appendix J
Weekly report regarding operational performance created by Operations Control (ZZ) to be used in weekly management team (MT) meeting of KLC.

Removed due to confidentiality

Appendix K
Performance conformance checklist used during turnarounds by the Duty Managers Ground Services. Used as input for the summary in appendix H.

Removed due to confidentiality

UNIVERSITEIT TWENTE. KLM

Appendix L
List of relevant KPIs as measured by KLC

Periodic & General

Name	Metric	Periodic?	Target?	Description
D0	Percentage	W, M, Q, Y	Y	% of flights in that period that leave on or before scheduled time
D15	Percentage	W, M, Q, Y	Y	% of flights in that period that leave within 15 minutes of scheduled time
A0	Percentage	W, M, Q, Y	Y	% of flights in that period that arrive on or before scheduled time
A15	Percentage	W, M, Q, Y	Y	% of flights in that period that arrive within 15 minutes of scheduled time
Completion factor	Percentage	W, M, Q, Y	Y	% of scheduled flights which actually have been flown
Registration changes	Number	W, M, Q, Y	N	Number of changes to which plane is assigned to a flight within 60 minutes of departure
Non-performance departments	Hours	W, M, Q, Y	N	Sum of delay caused by each department of KLC, identified using delay codes
Non-performance within department	Hours	W, M, Q, Y	N	Sum of delay caused by each activity within a given department, identified using delay codes

Ground Services

Name	Metric	Periodic?	Target?	Description
"Various services" ready at D-#	Boolean	M	Y	Whether the given service was finished at its designated moment, measured by hand for roughly 50 turnarounds per month =====>

Specific measurements with different service providers:

Toilet & Water

Name	Metric	Periodic?	Target?	Description
Throughput*	Percentage	W, (M?)	N	Percentage of total activities where planned time was equal to or bigger than actual time
Timeliness*	Percentage	W, (M?)	N	Percentage of total activities where start of activity was early enough to fully accommodate planned time before deadline
Average Duration*	Minutes, Seconds	W, (M?)	N	Average duration of all activities of that type in that period

Fueling:

Name	Metric	Periodic?	Target?	Description
# Delays caused by fueling**	Number	M, Y	Y	Count of delay codes assigned to fueling
# Refuels***	Number	M, Y	N	Count of total refuels in period
% Refuelings started before D-45*	Percentage	M, Y	Y	Percentage of refuelings started before D-45
# Defuels ***	Number	M, Y	Y	Count of total defuels (e.g. due to cancelled flight)
# Additional fuel ***	Number	M, Y	Y	Count of additional refuelings (e.g. due to changed weather conditions)
# Additional fuel after D-20*	Number	M, Y	N	Count of additional fuel starts after D-20
% Refuels with >100KG overshoot*	Percentage	M, Y	N	Percentage of refuels where the amount of loaded fuel was over 100 KG higher than requested

Flight Services

Name	Metric	Periodic?	Target?	Description
Flight Plan / Fuel	Hours	M, Y	N	Sum of delay caused around flight- & fuel plan approval, code 57, 61 & 62 (Flightplant N/A, - change, addition fuel required)
Co Crew	Hours	M, Y	N	Sum of delay caused by Cockpit crew, code 63,64,65, 69 (late crew boarding, crew missing, -special request, captain req security check)

Additional Department specific KPI's

Service	Deadline
Cleaning	D-21 (apron) D-20 (gate)
Water	D-15
Toilet	D-20 (F70) D-15 (E190)
Catering	D-23
Fueling	D-10
CDC	D-25
Crew on board	D-25
Meet&Greet TLO&Crew	D-25
Delivery of bags	D-30 (first) D-10 (last)
Arrival of busses (2 bus)	D-11 (first) D-6 (second)
Arrival of busses (3 bus, Dgate)	D-15 (first) D-6 (last)
Start boarding at Semi gates	D-20
Loadsheet on board (F70)	D-8
Talk out	D-4
Pushback in position	D-2
ADC & pax stairs removed	D-2

* Based on times as logged in CHIP
** Based on information from FIRDA
*** Based on information from FMS

All other KPI's used within KLC not mentioned here are either not quantitative or not relevant for the research / AP2K project

UNIVERSITEIT TWENTE.

Appendix M

List of measurements used as input for this research

#	Name	Type of measurement	Unit of measure	Description
General turnaround information				
A1	VisitDatabaseID	Descriptive		Unique identifier for a turnaround
A2	Date	Descriptive	Date	Date of departing flight
A3	AircraftRegistration	Descriptive		Callsign of involved aircraft
A4	AircraftType	Descriptive		Type of involved aircraft
A5	Station	Descriptive		Airfield at which the turnaround takes place
A6	InboundFlightNumber	Descriptive		Flight number of incoming flight
A7	InboundStation	Descriptive		Airport from which incoming flight departed
A8	InboundGateNumber	Descriptive		Gate at which arriving passengers have to deboard
A9	InboundLoadFactor	System generated	Percentage	Percentage of seats occupied on arriving flight
A10	OutboundFlightNumber	Descriptive		Flight number of outgoing flight
A11	OutboundStation	Descriptive		Airport to which outgoing flight is headed
A12	OutboundGateNumber	Descriptive		Gate at which departing passengers have to board
A13	LoadFactor	Calculated	Percentage	Percentage of seats occupied on departing flight based on aircraft capacity and passengers on board
A14	Weather	System generated	Descriptive	Type of weather during the hour of scheduled departure (Rain, Heavy rain, Snow, Thunder, Ice or Other)
General groundtime information				
B1	STA	System generated	Date + Time	Scheduled time of arrival (UTC)
B2	STD	System generated	Date + Time	Scheduled time of departure (UTC)
B3	ATA	System generated	Date + Time	Actual time of arrival (UTC)
B4	ATD	System generated	Date + Time	Actual time of departure (UTC)
B5	EBDE	System generated	Date + Time	
B6	ESTD	System generated	Date + Time	
B7	NumberofRegChanges	System generated	Integer	Total number of times the flight was assigned to a different aircraft
B8	LastestRegChange	System generated	Date + Time	Moment in time at which the last registration change took place
B9	RegistrationChangeInTime	Calculated	Boolean	Indicates whether the last registration change was before the required 1 hour of scheduled departure
B10	MinSTD	System generated	Date + Time	Earliest moment at which the aircraft can depart again, based on ATA and minimum ground time
B11	PlannedGroundTime	Calculated	Minutes	Timespan which the aircraft was scheduled to be at airport
B12	ReadGroundTime	Calculated	Minutes	Timespan which the aircraft was actually at airport
Delay information				
C1	Delay1	Manually generated	Descriptive	Primary assigned delay code
C2	Delayminutes1	Manually generated	Minutes	Number of minutes caused by first code
C3	Delay2	Manually generated	Descriptive	Secondary assigned delay code
C4	Delayminutes2	Manually generated	Minutes	Number of minutes caused by second code
C5	A0	Calculated	Minutes	Minutes between actual and scheduled time of arrival
C6	D0	Calculated	Minutes	Minutes between actual and scheduled time of departure
C7	ContributedD0	Calculated	Minutes	Total minutes of delay caused by this event / activity
C8	PropagatedD0	Calculated	Minutes	Minutes of delay propagated back to the previous event / activity
Specific events*				
D1	EventTime	System generated	Date + Time	Time at which the event was logged
D2	EventTardiness	System generated	Date + Time	Minutes between scheduled time of event in OCP-ZY20 and actual time
D3	EventDminus	System generated	Date + Time	Total minutes between actual arrival time and event time
D4	EventAplus	System generated	Date + Time	Total minutes between actual departure time and event time
Regular activities**				
E1	ActivityStart	Manually generated	Date + Time	Time at which the activity was started
E2	ActivityEnd	Manually generated	Date + Time	Time at which the activity was finished
E3	ActivityDuration	Calculated	Minutes	Total minutes between activity start and end time
E4	ActivityLateStart	Calculated	Minutes	Total minutes between scheduled time of activity start in OCP-ZY20 and actual start
E5	ActivityTardiness	Calculated	Minutes	Total minutes between scheduled time of activity completion in OCP-ZY20 and actual end
E6	ActivityDminus	Calculated	Minutes	Total minutes between actual arrival time and activity start
E7	ActivityAplus	Calculated	Minutes	Total minutes between acutal departure and activity end
Special activities				
F1	AxxArrStart	Manually generated	Date + Time	Axxicom bus collecting arriving PRM's & UM's start loading passengers at aircraft
F2	AxxArrEnd	Manually generated	Date + Time	Axxicom bus collecting arriving PRM's & UM's finished loading passengers at aircraft
F3	AxxDepGateStart	Manually generated	Date + Time	Axxicom bus delivering departing PRM's & UM's start loading passengers at gate
F4	AxxDepGateEnd	Manually generated	Date + Time	Axxicom bus delivering departing PRM's & UM's finished loading passengers at gate
F5	AxxDepOnlStart	Manually generated	Date + Time	Axxicom bus delivering departing PRM's & UM's finished offloading passengers at aircraft
F6	AxxDepOnlEnd	Manually generated	Date + Time	Axxicom bus delivering departing PRM's & UM's start offloading passengers at aircraft
F7	BusArrOnPos	Manually generated	Date + Time	First AAS bus collecting arriving passengers on position at aircraft
F8	BusArrStart	Manually generated	Date + Time	First AAS bus collecting arriving passengers loading passengers at aircraft
F9	BusArrEnd	Manually generated	Date + Time	Last AAS bus collecting arriving passengers finished loading passengers at aircraft
F10	BusDepOnlOnPos	Manually generated	Date + Time	First AAS bus delivering departing passengers on position at gate (ready to recieve passengers)
F11	BusDepOnlStart	Manually generated	Date + Time	First AAS bus delivering departing passengers start loading passengers at gate
F12	BusDepOnlEnd	Manually generated	Date + Time	Last AAS bus delivering departing passengers finished loading passengers at gate
F13	BusDepOflOnPos	Manually generated	Date + Time	First AAS bus delivering departing passengers on position at aircraft (ready to offload passengers)
F14	BusDepOflExStart	Manually generated	Date + Time	First AAS bus delivering departing passengers start offloading at aircraft
F15	BusDepOflEnd	Manually generated	Date + Time	Last AAS bus delivering departing passengers finished offloading at aircraft

* Events:	Description
CDO	Cargo door open
PDO	Passenger door open
ADO	All doors open
CDC	Cargo door closed
PDC	Passenger door closed
ADC	All doors closed
LoadSheetOB	Loadsheet on board
RTD	Ready to Depart

* Activities
TLO
Cargo
Cleaning
Crew Transport
Fuel
ExtraFuel
Pushback
Toilet
Water
Catering
Boarding pass scans

UNIVERSITEIT TWENTE.

KLM

Data availability of the different tasks

Activity	Turnarounds	Total Faulty	Missing	Out of Bounds	Short Duration	Extreme Duration
Bax *	69,901	1,954	642	1,081	105	214
Cleaning *	69,901	1,792	462	346	170	937
Crew Transport	69,901	17,260	1,408	960	1,931	13,632
Fuel *	69,901	2,234	407	956	114	874
Extra Fuel	69,901	67,626	67,574	41	15	-
Toilet *	69,901	2,240	597	521	852	429
Water	69,901	18,099	17,444	143	-	536
Catering *	69,901	11,468	7,187	15	1,459	2,811
PaxScan *	69,901	3,788	304	2,658	811	2,336
TLO	69,901	45,935	197	12,464	286	42,936
Pushback	69,901	26,827	24,069	154	2,580	80
Arrival Bus	69,901	53,202	51,635	276	1,308	2
Departure Bus	69,901	52,214	52,149	56	6	3
Arrival Axxicom	69,901	62,419	60,879	663	1,197	11
Departure Axxicom	69,901	65,214	64,141	21	6	1,060
Event						
CDO	69,901	31,167	31,127	40	N/A	N/A
CDC	69,901	28,101	28,089	12	N/A	N/A
PDO	69,901	28,294	28,251	43	N/A	N/A
PDC	69,901	28,015	28,005	10	N/A	N/A

UNIVERSITEIT TWENTE.

KLM