DELAY PROPAGATION AND ITS IMPLICATIONS IN AIRLINE OPERATIONS

by

Furkan Erdem

B.S., Mathematics, Boğaziçi University, 2018

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Industrial Engineering
Boğaziçi University
2021

ACKNOWLEDGEMENTS

I would like to express my special appreciation and thanks to my thesis advisor Prof. Taner Bilgiç for his endless support, encouragement, and patience. It has been a privilege to work with him. This thesis would not have been possible without his help.

I would like to thank to Prof. M. Selim Aktürk and Prof. Z. Caner Taşkın for their valuable comments and suggestions as members of my thesis committee.

I truly appreciate and value my parents, my sisters and, closest friends for managing the hardest part, by keeping me motivated and encouraged the whole time.

I would like to thank the Scientific and Technological Research Council Turkey for the financial support with the 2210-E National Scholarships Programme for Master's students.

ABSTRACT

DELAY PROPAGATION AND ITS IMPLICATIONS IN AIRLINE OPERATIONS

Airline companies try to increase their revenues, service level, and customer satisfaction in the competitive airline industry. Airline schedule planning is very crucial for airline companies to reach their objectives. While they are planning their schedules, they expect that there will be no disruption. On the other hand, there are plenty of incidences such as weather conditions, mechanical failure, traffic, and security issues that cause delays and disrupt the operations of an airline. Even though it is impossible to avoid the delay completely, there are ways to decrease the propagation of the delay. To cope with delay propagation, airlines put idle time between flights. However, idle time means inefficient use of aircraft resources. Thus, adjusting the idle time in the schedule is a critical task for planning departments in different aspects. In this study, flight time rescheduling and aircraft swapping are used to decrease the expected delay propagation. By applying these two options, the scheduled slack in the schedule is clustered at flights that are prone to propagate. It is aimed that the negative consequences of delay are reduced proactively while keeping total slack time in the schedule. Mathematical programming formulations the single layer model, the multi layer model, and the single layer model with swapping option are used. There is also an approximation for the single layer model with swapping option to overcome the computational issues. For the computational study, the flight schedule from Turkish Airlines is used. As a result, a decreased expected delay propagation is observed by rearranging the schedule without changing the cost of scheduled idle time.

ÖZET

RÖTAR YAYILIMI VE HAVAYOLU OPERASYONLARINA ETKİLERİ

Oldukça rekabetçi olan havayolu sektöründe bulunan havayolu şirketleri gelirlerini, hizmet düzeylerini, ve müşteri memnuniyeti seviyelerini arttırmaya çalışmaktadır. Tarife planlaması, havayolu şirketlerinin hedeflerine ulaşması için çok önemli bir operasyondur. Şirketler tarifelerini planlarken herhangi bir aksaklık olmayacağını beklerler. Öte yandan hava koşulları, uçaklardaki arızalar, trafik ve güvenlik sorunları gibi pek çok olay operasyonların aksamasına neden olur. Rötarı tamamen engellemek imkansız olsa da, rötarın yayılmasını azaltmanın yolları vardır. Havayolu şirketleri rötar yayılımıyla başa çıkmak için uçuşlarının arasına atıl süre koyar. Bununla birlikte, atıl süre, uçak kaynaklarının verimsiz kullanımı anlamına gelir. Bu nedenle, uçuş tarifesindeki atıl sürenin ayarlanması, planlama departmanları için oldukça önemli bir görevdir. Bu çalışmada, beklenen rötar yayılımını azaltmak için uçuş zamanını yeniden düzenleme ve uçak görev değişimi metodları kullanılmıştır. Bu iki seçenek uygulanarak, tarifedeki planlanan atıl süre, rötarı yaymaya eğilimli uçuşlarda kümelendirilir. Toplam atıl süre sabit tutularak rötarın olumsuz sonuçlarının proaktif olarak azaltılması tezin asıl amacıdır. Tek katmanlı model, çok katmanlı model ve uçak görev değişiminin mümkün olduğu tek katmanlı model problemi modellemek için kullanılmıştır. Ek olarak, uzun süren hesaplama süresinin yarattığı problemlerin üstesinden gelmek için, uçak görev değişiminin mümkün olduğu tek katmanlı modele alternatif sezgisel bir model oluşturulmuştur. Modellerin uygulanması için Türk Hava Yolları'nın uçuş tarifesi kullanılmıştır. Sonuç olarak, planlanan atıl sürenin maliyetini değiştirmeden uçuş tarifesini yeniden düzenleme yöntemiyle azalmış beklenen rötar yayılımı gözlemlenmiştir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

at_f	Scheduled arrival time of flight f
C	Set of all flight connections
$d_{f_i,f_j}^{m_i,m_j}$	Delay propagated from f_i which is flight of aircraft i which is experienced m_i minute of independent departure delay and f_j is flight of aircraft j with m_j minute of independent departure delay which is possible to swap with aircraft i
$d_{f,f'}^{m}$	Delay propagated from flight f which is experienced m minute of independent departure delay to f' where f and f' are connection flights
dt_f	Scheduled departure time of flight f
F	Set of daily flights
f'	Consecutive flight of flight f
k_f^+	Maximum limit for the change of departure time moved later
k_f^-	Maximum limit for the change of departure time moved earlier
M	Set of possible delay minutes
mtt	Minimum turnaround time of an aircraft
p_f^m	Probability of flight f experience m minute of independent departure delay
$r_{f_0}^m(f)$	Previous flight of f in the set $T_{f_0}^m$
s_{f_i,f_j}	Binary variable which equals one if aircraft of flight f_i swap with aircraft of flight f_j, zero otherwise
$T_{f_0}^m$	Set of possible flights that are affected by propagation of m minute independent delay at flight f_0
u_f	Idle time of flight f
x_f	Amount of change of scheduled departure time of flight f

LIST OF ACRONYMS/ABBREVIATIONS

ESE	Expected squared error
ISE	Integrated squared error
SLM	Single layer model
STA	Scheduled time of arrival
STD	Scheduled time of departure
MLM	Multi layer model

1. INTRODUCTION

The airline industry has a huge effect on the economy. According to IATA reports, revenues of airline companies are more than 876 billion US dollars in 2019 [1]. The total number of passengers carried by airlines is 4.5 billion and the total number of flights is 38.3 million in 2019 according to ICAO [2]. For this reason, there are plenty of decision-making problems for companies to improve their performance level. Delays are one of the most important disruptions which lead to customer dissatisfaction and loss of money. The total cost of the delay and additional time loss for airspace users and consumers is estimated as 33 billion dollars in 2019 for only United States [3]. Moreover, the demand for airline transportation is increased 4.2%, and airline capacity is increased by 3.4% in 2019 [4]. That imbalance between supply and demand may cause an increase in the cost of delay for airline companies. Thus, even a small increment in delay performance may save a large amount of money, and airline companies are looking for these improvements.

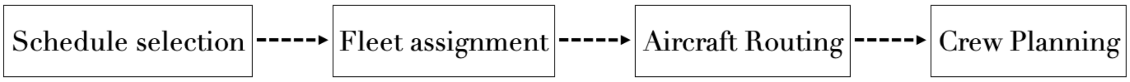


Figure 1.1. Airline schedule planning process.

Airline schedule planning is a process that starts almost one year before flights. In general, airline scheduling includes several steps that follow each other respectively. All these successive steps depend on each other. Schedule selection is constructed with historical data on passenger demand which assigns origin-destination and departure time. After the flight schedule is assigned, the next step is to assign the correct aircraft type to the predetermined flight. Aircraft routing is to determine aircraft that will operate a flight. A key point for aircraft routing is satisfying maintenance requirements for aircraft. The last step in airline schedule planning is a minimum cost set of crew itineraries or pairings that covers each flight exactly once.

Creating a model that will solve the problems of all these four different segments together may seem profitable and useful [5]. However, it is extremely complex and intractable, so the problem is required to be divided into pieces.

Airline companies make their plans and schedules by assuming no disruption will occur. However, in real life, there are plenty of incidences that lead to disruptions. Bad weather conditions, aircraft breakdown, crew delay, and congestion are the most significant factors which lead to disruption. In addition to cost, delays cause customer dissatisfaction that may affect negatively the future profitability of a company. To avoid these adverse conditions, airline companies try to catch the original plan as quickly as possible and make provisions against probable disruption causes. Delay propagation is a type of propagation that is an ongoing effect of previous flights. A delay propagation occurs when the previous flight has been delayed more than the slack time between two flights. Because the airline system operates as a closely interconnected network, it is subject to network effects, that is, a disruption in one place can quickly propagate to multiple other parts of the network [6]. Moreover, delay propagation is also a big problem for passengers since it means loss of connection flights. 40% of the total delays are caused by the late arrival of aircraft of the previous flight in the US [7]. Thus, minimizing delay propagation in the competitive airline industry is a crucial duty for the companies. Even though it is impossible to completely avoid delay, a key challenge is to minimize the ripple effect of delays, in other words, to make the schedule robust against disruptions.

There are various ways to cope with disruptions such as rearranging departure times to reallocate slack times, swapping aircraft, canceling a flight, calling reserve aircraft and crew. Disruptions in daily airline operations can be solved by trusting the experience of planning operators or some predefined procedures. There are also commercial tools that solve the disruption and scheduling problems of airline companies. In this study aircraft swapping and departure time change is used to cope with the effects of delay propagation. Reallocating the slack between flights is one way to absorb delay propagation which is caused by the late arrival of the previous aircraft.

The purpose of retiming is to make the slack time available where it is mostly needed operationally. Increasing slack means inefficient usage of a scarce source. In general, airlines design their schedules as tight as possible to use their aircraft resources efficiently, thus once disruption occurs it is possible that following flights will be affected. For this reason, allocating existing slack in the schedule efficiently may reduce the adverse effects of propagation. Swapping an aircraft of flights is another way to cope with delay propagation. In some cases, changing the planes of flights causes a more efficient distribution of idle times. Note that, both options explained above to cope with delay propagation are considered proactive that they try to decrease the negative consequences of delay propagation before a disruption occurs. In addition, total idle time will be the same after departure times are rescheduled and planes are swapped. In this way, the aim is to decrease the cost of delay propagation while keeping the cost of idle time as scheduled.

2. LITERATURE REVIEW

Since airline companies are trying to find ways to improve their efficiency in their operations, there are plenty of researches on different stages of the airline planning problem. There are more than 1000 papers about the aviation industry in the last 50 years [8]. In this chapter, we summarize some related works about our thesis.

2.1. Robust Planning - Flight Retiming Models

Airline companies make their plan by assuming no disruption will occur. However, in real life, they are confronted with significant expenses since disruption is frequent in aviation [9]. To minimize the effect of delay, planning departments take some precautions before a disruption occurs to make the schedule more robust. Holding standby crew and aircraft, adding extra slacks to the schedule, out and back type flights and increased cruise speed are some ways to increase the flexibility of the schedule which will help to catch the planned schedule in an easy way [10].

Stojkovic et al. [11] focus on the ways of recovering the schedule from minor disruptions such as headwinds or air traffic congestion. In their model, not only flight departure times are allowed to change but operational tasks are also free to hasten. For example, passenger boarding time of a flight can be shortened. They keep aircraft itineraries and crew rotations as scheduled. The objective function is the sum of the cost of shortened operations and the cost of passenger and crew inconvenience.

AhmadBeygi et al. [12] investigate how rearranging departure times between two connection flights can improve delay performance. The purpose of re-timing is to make the slack time available where it is most needed operationally. The earlier departure of a flight means increased slack time for the latter flight, but decreased slack time for the former flight. In the objective function, they product delay propagated from f_1 to f_2 where f_1 and f_2 are consecutive flights of same aircraft with the probability of experiencing m minutes of root delay of the flight f_1.

By weighing the connections with a probability of experiencing root delay, the model determines the optimal allocation of slack. While changing the departure time of a flight earlier or later does not change the total idle time of the whole system, but change their allocation. Thus, the total planned cost of the schedule remains the same. Note that, they consider both aircraft and crew as a scarce resource.

Lan et al. [9] define time windows that create time intervals for a departure time of flights. In their model, some flight copies are created in the time windows and the model chooses one of the flight copy which minimizes total disrupted passengers by sliding slack between flights. Note that the size of the problem depends on the width of the time windows. Narrower time windows, for instance, one minute, means more flight copies which increase the size of the problem. While model change flight departure times, they consider the problem in terms of time windows, flight schedule, fleet assignment, and aircraft routing decisions.

Chiraphadhanakul et al. [13] use flight re-timing with different objective functions. Minimizing total expected arrival delay and minimizing total expected propagated delay are two performance metrics in the paper. Beyond these two, they also introduce the notion of effective slack. Since propagated delay is zero until total arrival delay exceeds slack, a model whose performance metric is total propagated delay cannot distinguish the difference until threshold value which is the time that slack is equal to total arrival delay. By taking this notion into account, they define objective functions that maximize total expected aircraft connection slack and maximize total expected effective passenger connection slack. As expected, re-allocation of existing slack under different performance measure improves schedule performance even schedule adjustment is minor.

Aloulou et al. [14] increase the robustness while keeping schedule cost as it is with flight re-timing and re-routing. In their work, they define the robustness function and use it as a measure. Increasing the slack between two consecutive flight increase the robustness until a specific threshold and beyond this threshold, no improvement will occur in terms of robustness.

They calculate a flight-specific threshold according to empirical data about the delay, income, and passengers' connections about that flight leg. As a result, they found that both propagation effect of disruption, the total number of a passenger who miss their connection flight and total robustness is improved.

Desphande et al. [15] focus the effect of scheduled block time allocation on airline on-time arrival performance of an airline. They estimate the probability of scheduled on-time arrival of a flight and then they try to understand the underlying mechanism of scheduled on-time arrival probability. According to Desphande et al. [15] on-time performance of an airline is determined by four main factors, namely, operational, competition, cost minimization, and revenue maximization. They found from their data, airlines put shorter planned block time than actually required time to operate a flight. Multivariate regression is used for related estimations in the paper. The main drivers of the overage to underage cost ratio are analyzed.

In this thesis, we cope with expected disruptions by flight re-timing and aircraft swapping. By using flight departure re-timing, we allow an aircraft to depart earlier or later than its original departure time. The model will determine the amount of change of departure times according to propagation tendency of each flight individually. Aircraft swapping is another tool that makes the schedule more robust against disruption. In our model, the objective is expected propagated departure delay and all calculations are made using probability values which are calculated from past flight data.

2.2. Airline Recovery Problems

Planning departments have to solve the negative consequences of disruption as quickly as possible when the schedule is disrupted. Aircraft recovery problem seeks the best possible solution to straighten the schedule by saving aircraft from the disrupted situation.

Rosenberger et al. [16] use aircraft rerouting to cope with disruptions. While they are swapping aircraft they consider both airline's and FAA's regulations and maintenance requirements. They consider aircraft recovery problems from two sides which are aircraft disruption and station disruption. Aircraft disruption is mainly caused by a mechanical failure in aircraft or prolonged in-flight delays. On the other hand, station disruption is a result of bad weather and congestion at the airport. The model in the paper is a sort of set packing problem where the flight leg is either stayed as planned or canceled. They use a heuristic method to select aircraft that are available for rerouting. By using aircraft selection heuristic they deal with the hassle of set-packing problems. In the objective function, all costs of rerouting and cancellation are included.

Aktürk et al. [17] make flight time rescheduling with cruise speed control. They change the scheduled flight times by increasing (or decreasing) cruising speed. Their model tries to minimize the total cost of recovery actions and the cost of delay. There is a trade-off between increasing cruise speed and flight delay. Higher cruise speed means more fuel burned which has negative environmental effects and cost of extra fuel burned. Another recovery action in their model is aircraft swapping which incurs crew deadhead cost. They use a model that does not track an aircraft individually which makes the problem size substantially smaller. The objective function in their model includes all costs of fuel, CO_2 emission, delay an aircraft swap. They use conic mixed-integer programming to cope with computational complexity of the problem.

Duran et al. [18] try to increase the robustness of the flight schedule by controlling cruise speed. They assume that cruise time is controllable by changing cruise speed, whereas non-cruise time is subject to uncertainty. Their objective is to minimize idle time and fuel costs. While minimizing the cost of idle time and fuel, they assure that minimum passenger connection service level is provided flight departure time change or cruise speed adjustment. They found that %60 percent decrease in the cost of idle time can be provided by increasing the fuel cost %2 percent.

Vos et al. [19] use an innovative dynamic modeling framework to solve the aircraft recovery problem. The traditional approach solves the problem by assuming that disruptions are known at the starting point. In their work, the model is solved when disruption occurs. In this way, they solve the problem by having the same information as the airline planning department. The model also changes the previous recovery decisions within a certain time-window if it will improve the objective function. Objective function measures the direct cost of delay and the cost of missed opportunity cost because of disrupted resources. To solve the problem at a reasonable time they use the aircraft selection algorithm.

In this thesis, the main purpose is to recover disruption and its negative effects before it occurs. It may be more difficult and expensive to solve the large scale scheduling problem of an airline actively. It is also important to solve the problem as fast as possible since late decisions can cause more problems. By looking past flight data, our model calculates possible flights that cause propagation and tries to prevent them proactively.

2.3. Probability Estimation

AhmadBeygi et al. [12] use the original one-year length schedule of airline and generate an empirical distribution based on the real flights according to the origin airport of flight. At first, they remove the propagation factor from each delay to get the root delay factor. Then, they use the probability mass function of root delay for both the objective function and simulation.

Lan et al. [9] assume that an aircraft can arrive destination point earlier, on time, or late. They assert that possible candidates for arrival delays are gamma distribution, log-normal distribution, and Weibull distribution. By using chi-square statistics and Kolmogorov test, they found that 84 percent of data follow a log-normal distribution with 0.01 significance level.

Note that, their estimation does not care about departure delay, instead they directly find arrival delay probabilities. In our model, we assume that m minutes departure delay is carried as m minutes arrival delay. That means, experienced delay at departure cannot be shortened (stretched) at the cruising stage.

Coi et al. [20] search the probability mass function of delay by dividing them as factors. One for the propagation factor and one for non propagation factor. They found that delay that is affected by non-propagation factors exhibits power-law distribution. On the other hand, a delay that includes propagation follows shifted power-law distribution. Note that, by doing these they found a general distribution function of delay for the flights. It should be noted that it is a general case (non-parametric) which do not take care external events such as bad weather, airport closure etc.

Belcatro et al. [21] make a prediction for flight delays based on weather conditions. They set a threshold value and then use random forest to test whether a given flight will be above or below the threshold. For 15 minutes threshold, they achieve 74.2 percent accuracy. They found that even without the weather data model achieve 69.1 percent accuracy which means there exists an underlying pattern of delay.

Choi et al. [22] create a model by using different machine learning and data mining methods to predict airline delays. They fed the model with the original schedule and weather forecast and the model makes binary classification prediction about on-time performance of the flight. Decision trees, random forest, the AdaBoost, and the kNearest-Neighbors are the different techniques used in prediction. They found that random forest has the best performance to predict on-time performance of the airline.

Ye et al. [23] improve a model that predicts aggregate flight delay within a one-hour forecast horizon with the help of supervised learning methods. Flight schedule information and meteorological information is used to make a prediction.

Multiple linear regression, support vector machine, extremely randomized trees, and gradient boosting methods are selected for prediction modeling. They found that weather is not the most important factor to make a prediction. Light GBM gives the best results at the forecast horizon.

Hansen et al. [24] analyzed the factors that affect delay. Time of the day, traffic volume, weather conditions, seasonality, and secular effects are selected factors that may possibly affect the delay in the paper. They emphasize that delay is caused by some factors and their compound effects. For instance, a more traffic volume in the morning may cause more total daily delay since the higher traffic volume in the morning causes delay propagation to the subsequent flights of the same day. Their model explained approximately %70 of the variation in average delay.

In our work, we predict the probability of experiencing root delay of each flight by considering its departure time features such as meteorological conditions, departure time, schedule related information, etc. Multivariate kernel density estimation is used to estimate the delay probability of each flight.

3. PREDICTING DELAYS

At the first step, the probability distribution of delay should be calculated for each flight to use in the model. There exist an actual and planned flight schedule from Turkish Airlines. From the difference between realized and planned schedules, independent delay can be calculated. By using the weather data and other flight schedule information, the relationship between independent delay and factors can be founded.

3.1. Data

This thesis is based on the data provided by Turkish Airlines. Data consists of 74707 flights of 307 different aircraft and planned and realized schedule for three months (November 2019, December 2019, and January of 2020). Each flight has Istanbul Airport as either arrival or departure station since it is a hub for the airline. Scheduled departure time (SDT) and scheduled arrival time (SAT) are planned departure and arrival time of a flight. From SDT and SAT, expected in-flight time and slack time can be calculated. Hours of days are divided into four as six hours intervals for convenience. The registration code of an aircraft that is assigned to the flight is known, thus it is possible to track the origin and destination station of the aircraft. For each flight, data includes various characteristic information about that flight. An example from the data is given in Table 3.1.

Table 3.1. Sample Data.

Date	Traffic	Registration	Departure	Arrival	SDT	SAT	Cloudiness	Wind (m/s)	Temperature (C)
22.01.2020	Domestic	TCJHE	MQM	IST	07:25	09:55	Few	4	13
22.01.2020	International	TCJVK	ECN	IST	07:25	09:20	Scattered	7	18
22.01.2020	Domestic	TCJRK	IST	TZX	07:25	09:20	Clean	4	9
22.01.2020	Domestic	TCJHO	GNY	IST	07:30	09:50	Scattered	6	12
22.01.2020	International	TCJSU	KIV	IST	07:35	09:15	Clean	11	5

In addition to SDT and SAT, actual departure and arrival times are also known for each flight thus related values for departure delay can be calculated by subtracting actual time from the scheduled time. Independent delay is found by subtracting departure delay from slack. To find an independent departure delay of a flight, departure delay is subtracted from the slack of related flight. Note that, slack of flight is assumed as 35 minutes for all flights. In this thesis, we will not add a delay of cockpit and cabin related delay since our data does not include them.

Since the weather is one of the most important factors that cause 75 percent of delays [25], weather conditions related to flights are matched [26]. The weather conditions around the airport are called meteorological aerodrome report which is used by airline companies. Temperature, wind, humidity, cloudiness, visibility, and dew point some of the information included in the meteorological aerodrome report. In this study, a meteorological aerodrome report is used for each flight as weather data. A more detailed sample of data is given in the Appendix.

3.2. Kernel Density Estimation

Kernel density estimation is a useful tool to find the probability density function of a random variable. General form of kernel density estimation where $x_1, x_2, x_3, ...x_n$ are independent and identically distributed sample with size n from a population and h is the bandwith :

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{n=1}^{\infty} K(\frac{x - x_i}{h}) \tag{3.1}$$

where K is characteristic of Kernel function. It can be chosen as Gaussian, uniform, Epanechnikov, triangular [27]. Symmetric kernel functions have the following properties:

- $K(u) = K(-u)$
- $\int K(u)\, du = 1$

- $\int uK(u)\,du = 0$
- $\int u^2K(u)\,du = k \neq 0$

Further investigation about kernel density estimate requires its error calculation. More details can be found in the third and fourth chapter of Silverman's book [27] and fourteenth chapter of Shalizi's lecture notes [28]. Firstly, the expected value and variance of the estimate can be calculated as

$$\mathbb{E}[\hat{f}_h(x)] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[\frac{1}{h}K(\frac{x-x_i}{h})] \tag{3.2}$$

$$= f(x) + \frac{h^2 f''(x)}{2}\int K(u)u^2\,du + o(h^2). \tag{3.3}$$

The definition of the "o" encountered at (3.3) can be found in Section 7.6 in [29]. Assume that $\int K(u)u^2\,du = \sigma_K^2$. Then bias of estimation is

$$\mathbb{E}[\hat{f}_h(x)] - f(x) = \frac{h^2\sigma_K^2 f''(x)}{2} + o(h^2). \tag{3.4}$$

It means narrower bandwith (small h) makes estimation bias lower. Variance of kernel density estimation is

$$\mathbb{V}[\hat{f}_h(x)] = \frac{1}{n}\mathbb{V}[\frac{1}{h}K(\frac{x-X}{h})] \tag{3.5}$$

$$= \frac{f(x)}{nh}\int K^2(u)\,du + O(1/n). \tag{3.6}$$

The definition of the "O" encountered at (3.6) can be found in Section 7.6 in [29]. From the equation above, variance goes to zero when $nh \to \infty$ while sample size $n \to \infty$. As a conclusion, bandwith should go to zero but its rate should be smaller than $1/n$.

Then by the equality:

$$ESE = [\mathbb{E}[\hat{f}_h(x)] - f(x)]^2 + \mathbb{V}[\hat{f}_h(x)] \tag{3.7}$$

$$= \frac{h^4 \sigma_K^4 (f''(x))^2}{4} + \frac{f(x)}{nh} \int K^2(u)\, du + c \tag{3.8}$$

$$ISE = \frac{h^4 \sigma_K^4}{4} \int (f''(x))^2\, dx + \frac{\int K^2(u)\, du}{nh}. \tag{3.9}$$

Where ESE stands for expected squared error, ISE stands for integrated squared error. Taking the derivative with respect to h and setting it to zero to find optimal bandwith gives

$$h_{opt}^3 \sigma_K^4 \int (f''(x))^2\, dx - \frac{\int K^2(u)\, du}{n h_{opt}^2} = 0 \tag{3.10}$$

$$h_{opt}^3 \sigma_K^4 \int (f''(x))^2\, dx = \frac{\int K^2(u)\, du}{n h_{opt}^2} \tag{3.11}$$

$$h_{opt} = \Big(\frac{\int K^2(u)\, du}{\sigma_K^4 \int (f''(x))^2\, dx}\Big)^{1/5} n^{-1/5} \tag{3.12}$$

$$= O(n^{-1/5}). \tag{3.13}$$

When $h_{optimal}$ inserted into the equation of ISE, we get $ISE = O(n^{-4/5})$. Note that, estimation error with known parametric form has $ISE = O(n^{-1})$. Thus by having large enough data and finding bandwith by cross-validation may give near-optimal estimation by the law of large numbers.

Kernel density estimation can be done for multi dimensional variables. Where the general form is

$$\hat{f}_h(x) = \frac{1}{n \prod_{i=1}^{p} h_j} \sum_{i=1}^{n} \prod_{i=1}^{d} h_j K_j\Big(\frac{x^j - x_i^j}{h_j}\Big). \tag{3.14}$$

By applying similar error calculations as one-dimensional data, expected integrated squared error is $O(n^{-4/(4+p)})$ where p stands for dimension. For higher dimensions, the asymptotic rate is drastically low. Thus, estimating with less dimension is crucial for both the simplicity and reliability of estimation. Finding the best bandwith by cross-validation also requires a long computation time. Moreover, computation time increases exponentially with increased sample size.

Since multi dimensional kernel density estimation is defined, conditional kernel density also can be written by formula

$$\hat{f}_{Y|X}(y|x) = \frac{\hat{f}_{X|Y}(x|y)}{\hat{f}_X(x)} \tag{3.15}$$

where bandwith for each dimension namely h_x and h_y are founded to minimize expected integrated squared error.

3.3. Application to Delay Distribution Estimation

In this thesis, probability distributions of delay under given conditions are required to calculate expected delay propagation. Multiple conditions affect the delay such as wind, temperature, cloudiness, departure time, traffic, route, and so on. However, for tractability of multivariate kernel density estimation number of factors should be restricted. Wind, cloudiness, traffic, and the departure time interval are factors that are selected for the estimation problem. The individual effect of each selected factor can be seen in Figures 3.1, 3.2, 3.3 and 3.4.

Flights that experience delay higher than 90 minutes are excluded from the data set. There are only 671 flights with a delay of more than 90 minutes. Airline schedule data is matched with only weather data of Istanbul departure flights. Thus, half of the flights have weather data. In the final, 32832 flights remain on hand.

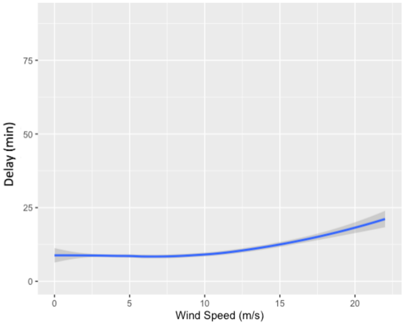


Figure 3.1. Wind-Delay Length.

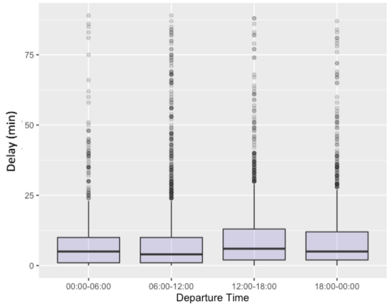


Figure 3.2. Departure Time-Delay Length.

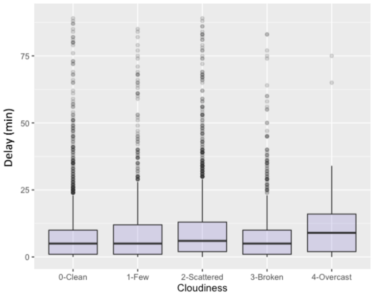


Figure 3.3. Cloudiness-Delay Length.

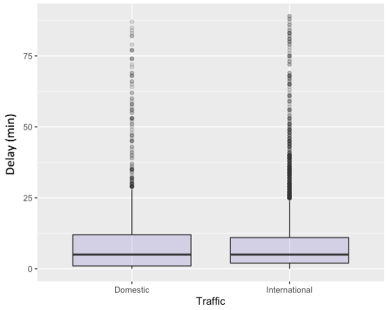


Figure 3.4. Traffic-Delay Length.

From these data, two samples are constructed. The first sample is a weighted sample that includes all flights with extreme weather conditions. And the second sample is constructed randomly. Since the estimation is computationally demanding, the sample size is restricted to 15000 flights. The ‘np’ [30] package in R programming language is used to find conditional multivariate kernel density estimation. In the ‘np’ package, the method for selecting bandwidth is selected as cross validation. Computation time is 21.4 hours for the first sample, 21.2 hours for the second sample. The codes in this thesis are calculated on Intel® Core® i5 1.6 GHz CPU.

The estimation has a multilevel response variable, namely the probability of a delay minute. However our data does not include this multilevel response, instead, it just has an actual departure delay of a flight under specific conditions. Thus, it is not possible to calculate error directly from the data.

Instead of calculating the error, we try to keep the model simple by restricting the number of factors and solving the estimation problem with the larger sample as far as the computational time allows. By taking these precautions and ISE performance of multivariate kernel density estimation, we expect to get acceptable estimation results.

Since departure delay is recorded as an integer in the data, after finding related density functions of flights integrating the function is necessary. To solve this issue, at first cumulative conditional multidimensional kernel density functions are utilized. Then the equation to find probability values of m minutes independent delay under given conditions is

$$P(m|X = x_i) = CDF(m|X = x_i) - CDF(m - 1|X = x_i) \tag{3.16}$$

where X stands for conditions namely, wind speed, traffic, cloudiness, and departure time interval.

Table 3.2. Different condition levels that determine probability distributions.

Condition	Traffic	Departure Time Interval	Cloudiness	Wind (m/s)
Condition 1	Domestic	00:00-06:00	Clean	0
Condition 2	Domestic	12:00-18:00	Few	7
Condition 3	International	12:00-18:00	Scattered	11
Condition 4	International	18:00-24:00	Broken	20

In Table 3.2 four different conditions and their related levels are stated to explain how conditional kernel density estimation will be used. The first condition in Table 3.2 is the most favorable to depart on time from an airport. Whereas, the fourth condition can be seen as the worst for departure. Note that, the probability distribution for all possible combinations (920 different combinations) of these four different factors are calculated. Estimated probabilities of delays under given conditions can be seen in Figures 3.5 and 3.6. The weighted and randomized samples show similar results.

As expected fourth condition has the highest expected independent departure delay value, whereas the first condition has the lowest expected independent departure delay value for both weighted and randomized samples. However, since the weighted sample includes more data that are relatively severe in terms of weather, condition 4 shows the difference in the two sample sizes, especially at higher delay minutes. In this thesis, both weighted and randomized samples will be used to see how sample selection affects the independent delay probability distribution. For instance, the probability estimation of independent delay for higher minute values is more realistic in estimation with the weighted sample as it can be seen from Figure 3.5.

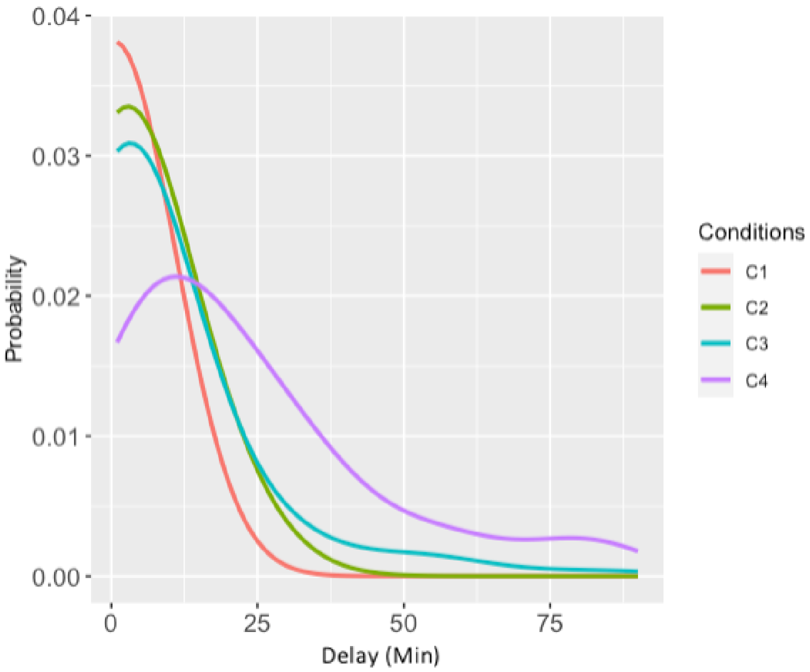


Figure 3.5. Result of Kernel Estimation for Flight Delay with Weighted Sample.

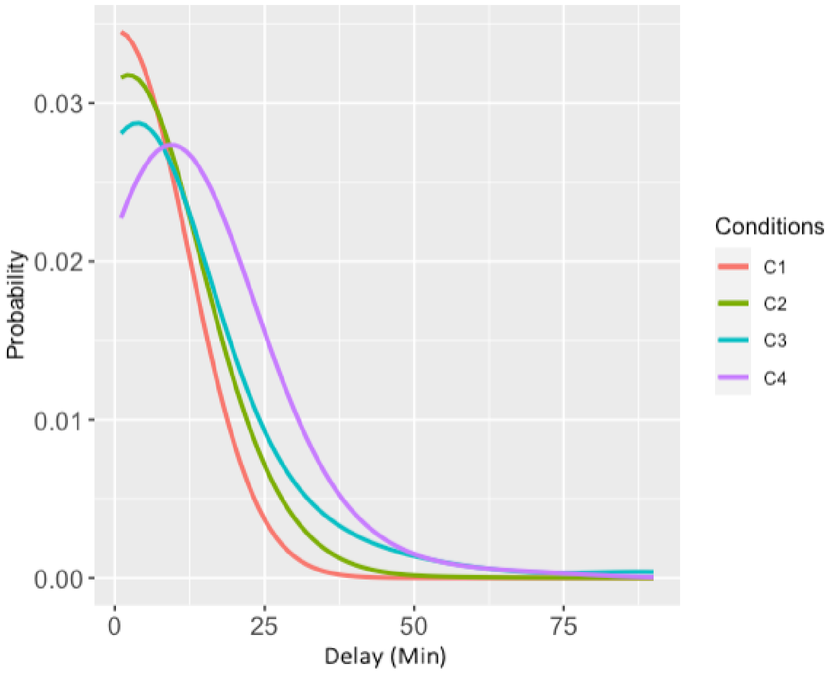


Figure 3.6. Result of Kernel Estimation for Flight Delay with Randomized Sample.

4. DELAY PROPAGATION MODELS

In this chapter, at first, some related definitions and concepts are given about airline scheduling. The objective is a minimization of expected delay propagation of an aircraft on a daily basis. Two main actions exist to decrease delay propagation namely flight re-timing and aircraft swapping. There are also two different ways to calculate expected delay propagation, namely single layer assumption and multi layer assumption. Based on these different options and assumptions, the improvement on expected delay propagation is aimed.

4.1. Related Definitions

Delay can be defined as the difference between actual departure (arrival) time and scheduled departure (arrival) time of aircraft at its flight [31]. It has two sides namely, departure delay and arrival delay.

Independent delay is a type of delay which has an external reason such as bad weather, mechanical failure, congestion at an airport. Independent delay is inevitable by its nature. For instance, there is no operational solution for the problems like excessive wind or mechanical issues of an aircraft.

Propagation delay is a type of delay which results from the late arrival of an aircraft. Delayed crew or delayed aircraft from previous flights lead to disruption on the next flights and cause delay propagation. Note that total departure delay includes independent departure delay and propagated delay from previous flight.

Minimum turnaround time is a time that refers minimum time required to prepare flights for the next flight. Refueling, catering, cleaning are some operational tasks that are done at turnaround time. It is generally between 25 and 45 minutes that depends on the operator, aircraft, airport, etc.

Slack time refers to an idle time of an aircraft where no operation is done about flight or airplane. It is placed into the airline schedule to absorb unexpected disruptions. If there is no slack between flights of an aircraft, then that flight becomes prone to propagate delay to subsequent flights. On the other hand, longer idle times mean inefficient use of aircraft resources.

These concepts can be connected by mathematical equations as follows:

$$Slack_{f_1} = Scheduled\ Departure\ time_{f_2} - Scheduled\ Arrival\ time_{f_1} - MTT$$
$$Propagated\ Delay_{f_1,f_2} = maximum(0, Independent\ Departure\ delay_{f_1} - Slack_{f_1})$$

where f_1 and f_2 are connection flights and MTT stands for minimum turnaround time.

In Figure 4.1 the idea of delay propagation is explained. Continuous lines stand for scheduled departure and arrival times and dashed lines for actual departures and arrival times. In the figure, the flight f_1 experiences m minutes of independent delay and there is no independent delay for f_2. Propagation occurs from flight f_1 to f_2 since m is larger than idle time of flight f_1. As a result, flight f_2 experiences $p_{1,2} = m - slack_{f_1}$ minutes of propagated delay.

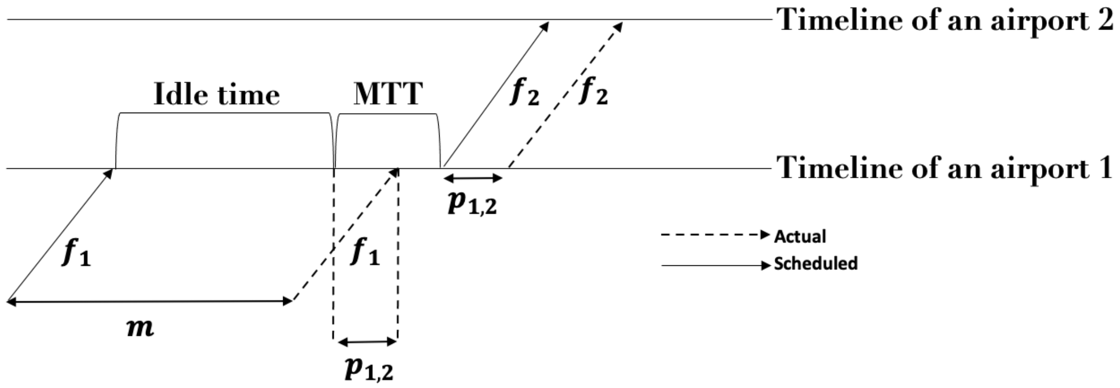


Figure 4.1. Idea of delay propagation.

Flight retiming is changing the departure time of a flight. The departure time of a flight can be either moved later or earlier from its original time. By moving departure time of an aircraft later slack of its previous flight increases. The earlier departure time of an aircraft increases the slack of its connection flight. AhmadBeygi et al. explain the idea of flight retiming in their paper [12].

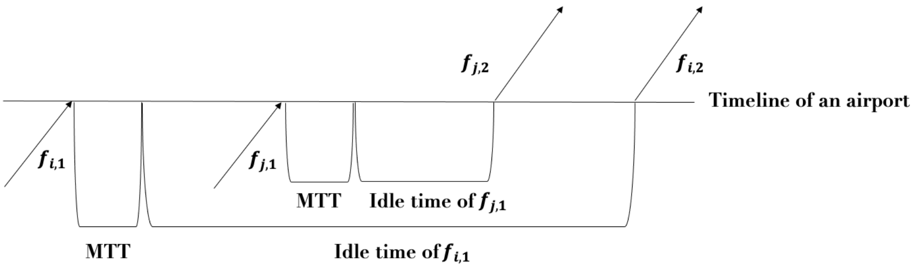


Figure 4.2. The schedule before aircraft swapping.

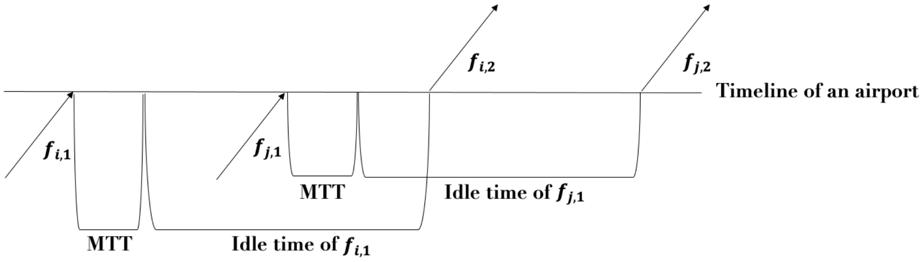


Figure 4.3. The schedule after aircraft swapping.

Aircraft swapping is an operation that aircraft of two different flights swap with each other. By swapping operation not only aircraft but also the crew swap with each other. Its main purpose is distributing total idle time evenly between flights. Evenly distributed idle time means balancing the robustness of flights. More clearly, swapping increases the robustness of flights that have shorter idle time [32]. Distributing the idle time between the daily flights homogeneously results in the homogeneous distribution of expected delay propagation. That is beneficial for the economy of the airlines since the cost of delay is nonlinear [33]. It is more expensive if an aircraft experiences a 30 minutes delay rather than each of two aircraft that experience 15 minutes delay. In Figure 4.2, there is the schedule before the swapping occurs.

Aircraft i has redundant idle time at its first flight. On the other hand, aircraft j has a relatively short idle time which makes it prone to propagation. In Figure 4.3, aircraft i and j swap their duties, namely $f_{i,2}$ and $f_{j,2}$. Even total idle time in the schedule is the same as before swapping occurs, they are homogeneously distributed between aircraft. Consequently total robustness of the schedule increases.

4.2. Description of the Delay Propagation Model

The main purpose of this model is to decrease the effects of probable delay propagation proactively by minimizing the total expected propagated delay. At first, the planning department of the airline calculates relative probabilities of delay of flights in the morning by looking at the weather forecast of the day and some flight related information. After calculating delay probabilities of flights, the model changes the departure times of the flights or swap the planes of the flights to redistribute the idle times between flights where it is mostly needed according to propagation tendency. Another decision is swapping duties of aircraft to distribute slack homogeneously. The advantage of this model is that total idle time does not change after flight time redistribution and aircraft swapping, so there is no additional schedule cost.

There are some assumptions used in the model as follows:

- Block time is fixed, which means 10 minutes delay at departure airport results as 10 minutes delay at the arrival airport. The word delay refers to departure delay throughout the thesis.
- After departure times of flights are changed and planes are swapped, demand profile, cabin-cockpit crew duty time restrictions, and customer satisfaction are preserved.
- Minimum turnaround time is 35 minutes for all narrow body aircraft that are used in the model.
- Aircraft maintenance operations are done at night so that they do not intervene the operations.

- The departure time of first and last flight of the day is not moved earlier or later to keep total idle time the same as before rescheduling.
- There is no propagation from the last flight of the day to the first flight of the next day.
- All narrow body planes used in the model have same physical features such as range and seat capacity to avoid passenger spill when swapping occurs.

4.3. Independent Delay Probabilities of Flights

Since models used in this thesis calculate expected values of delay propagation, related independent delay probabilities are required. As explained in Chapter 3, related probabilities are calculated by using multivariate conditional kernel density estimation. There are four different factors that affect the result of independent delay probability distribution of related flights, namely average wind speed, cloudiness, departure time interval, and traffic. All independent delay probability distributions of possible combinations of these four different factors are calculated by using the past flight schedule. For given conditions of the flight f, independent probability distribution of the flight is calculated exhaustively for all conditions using past data. Thus, p_f^m is a function of conditions of flight f. A more precise notation would be $p_f^m(x)$ as a function of a feature set X, where we use X={Wind, Cloudiness, Departure Time, Traffic}. Therefore, for planning purposes, future values of X must be determined. Note that weather information is known while calculating probabilities of flights by using multivariate conditional kernel density estimation. In reality, the planning department calculates independent delay probability distribution in the morning by using the daily weather forecast. Thus, it is assumed that that the daily weather forecast is accurate enough. Another feature that affects the independent delay probability distribution is the departure time of the flight. The model we discuss in this chapter allows for changing the departure time. But this change is limited. Therefore, it is assumed that changing the departure time of flight does not affect the probability distribution of flight.

4.4. Single Layer Model

The single layer model is the basic model that assumes propagation of a flight affects only its connection flight not the later flights of the day. The model is a similar model to the AhmadBeygi et al [12] except the way of calculation of probabilities and the crew propagation effect.

Table 4.1. Sets, variables and parameters.

Indices	
M	Set of possible delay minutes, m=1,2,3,...
C	Set of all flight connections
F	Set of daily flights, f=1,2,3,...
Decision Variables	
$d^m_{f,f'}$	Delay propagated from flight f which is experienced m minute of independent departure delay to f' where f and f' are connection flights
x_f	Amount of change of scheduled departure time of flight f
u_f	Idle time of flight f
Parameters	
mtt	Minimum turnaround time of an aircraft
dt_f	Scheduled departure time of flight f
at_f	Scheduled arrival time of flight f
k^+_f	Maximum limit for the change of departure time moved later
k^-_f	Maximum limit for the change of departure time moved earlier
p^m_f	Probability of flight f experience m minute of independent departure delay
f'	Consecutive flight of flight f

The single layer model can be stated as:

$$\text{Min} \sum_{m \in M} \sum_{(f,f') \in C} p_f^m d_{f,f'}^m \tag{4.1}$$

subject to

$$\begin{aligned}
& u_f = dt_{f'} - at_f - x_f + x_{f'} - mtt && \forall (f,f') \in C && (4.2)\\
& d_{f,f'}^m \geq m - u_f && \forall m \in M, \forall (f,f') \in C && (4.3)\\
& -k_f^- \leq x_f \leq k_f^+ && \forall f \in F && (4.4)\\
& u_f \geq 0 && \forall f \in F && (4.5)\\
& d_{f,f'}^m \geq 0 && \forall m \in M, \forall (f,f') \in C && (4.6)
\end{aligned}$$

Decision variables u_f, x_f, $d_{f,f'}^m$ are all continuous type. The resulting model is LP with a special feature that the solution for the model gives an integer x_f which is explained in AhmadBeygi et al [12] . In the objective function (4.1), the parameter p_f^m is flight specific and calculated by conditional multivariate kernel density estimation (see Chapter 3). The variable $d_{f,f'}^m$ stands for the propagation of m minute of independent delay from flight f to its connection flight f'. By summing this product over all possible minutes of delay and connection flights we get the expected value of delay propagation of an aircraft on its all flights in the day. The constraint (4.2) is the definition of slack of flight f. (4.3) is the definition of propagated delay from flight f to its connection flight f'. The constraint (4.4) limits the change of flight departure time. (4.5) and (4.6) are nonnegativity constraints for slack and delay propagation.

4.5. Multi Layer Model

Multi layer model has the same idea as the single layer model except that it assumes the ongoing effect of propagation continues until it is fully absorbed by slack between flights. This model is similar to the model in AhmadBeygi et al [12]. The multi layer model is more realistic than the single layer model.

The propagation of each flight is calculated by assuming that propagated delay from the previous flight does not include the independent delay of the current flight. In Figure 4.4, the daily flights of an imaginary aircraft are presented as nodes. The idle time of the first flight is 15 minutes and the idle time of the second flight is 20 minutes. Assume that the first flight experiences 40 minutes of independent delay. Then there will be 25 minutes of delay propagation to the second flight. Since independent delay at the first flight is not absorbed by the second flight, it will propagate to the third flight. There will be 5 minutes of propagation from the first flight to the third flight. The independent delay at the second flight is not added to the 25 minutes of propagated delay from the first flight while calculating the delay propagated to the third flight. Thus, the multi layer model calculated the total daily delay propagation of an aircraft by creating an independent delay at the first flight of the day individually and track that independent delay until it is fully absorbed while assuming there is no other independent delay at the subsequent flights. Then, the multi layer model repeats these processes for each of the subsequent flights. The model does not include independent delay at subsequent flights to the propagated delay from the previous flights since it makes the modeling process extremely difficult. Therefore this model may underestimate the propagated delay.

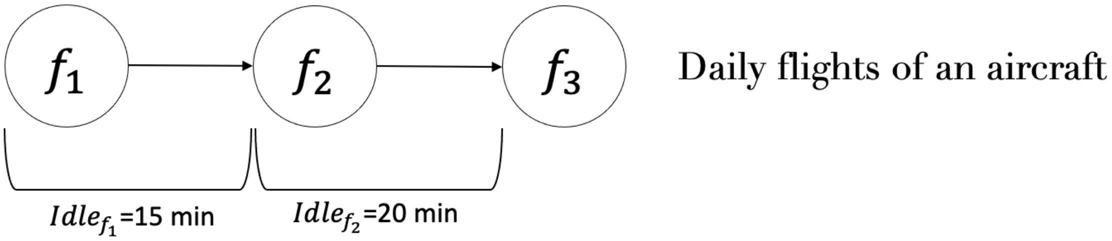


Figure 4.4. Idea of multi layer model.

Two new notations are required for the multi layer model.

- $T^m_{f_0}$= Set of possible flights that are affected by propagation of m minute independent delay at flight f_0
- $r^m_{f_0}(f)$= Previous flight of f in the set $T^m_{f_0}$

The multi layer model can be stated as:

$$\text{Min} \sum_{m \in M} \sum_{(f_0 \in F, f \in T^m_{f_0})} p^m_{f_0} d^m_{f_0,f} \tag{4.7}$$

subject to

$$\begin{aligned}
& u_f = dt_{f'} - at_f - x_f + x_{f'} - mtt && \forall (f, f') \in C && (4.8)\\
& d^m_{f_0,f} \geq m - u_{f_0} && \forall m \in M, \forall (f_0 \in F, f \in T^m_{f_0} : r^m_{f_0}(f) = f_0) && (4.9)\\
& d^m_{f_0,f} \geq d^m_{f_0,r^m_{f_0}(f)} - u_{r^m_{f_0}(f)} && \forall m \in M, \forall (f_0 \in F, f \in T^m_{f_0} : r^m_{f_0}(f) \neq f_0) && (4.10)\\
& d^m_{f_0,f} \geq 0 && \forall m \in M, \forall (f_0 \in F, f \in T^m_{f_0}) && (4.11)\\
& -k^-_f \leq x_f \leq k^+_f && \forall f \in F && (4.12)\\
& u_{f_1} \geq 0 && \forall (f_1, f_2) \in C && (4.13)
\end{aligned}$$

Similar to the single layer model, multi layer model is also LP model that gives integer x_f values because of its special characteristics [12]. The difference of the multi layer model from the single layer model is caused by equation (4.10) which states that if m minute of independent delay is not absorbed by its direct connection flight then propagation continues through subsequent flights by subtracting residual propagation from idle time of the related flight.

4.6. Single Layer Model with Swapping Option

Single layer model with swapping option is just an extension of single layer model by adding the swapping option. Note that even objective functions look different for two models, when swapping option is avoided at the extension model, two models give exactly the same results. Thus, it is possible to compare the objection function values directly to see the effect of departure time changing and aircraft swapping options. Note that, sets, parameters, and variables are defined again for the swapping models since the notation is slightly different from the model without swapping.

Table 4.2. Sets, variables and parameters.

Indices	
M	Set of possible delay minutes, m_i=1,2,3,...
A	Set of aircraft, i=1,2,3,...
S_i	Set of aircraft that can be swap with aircraft i, j=1,2,3,...
F_i	Set of daily flights of aircraft i, f_i=1,2,3,...
Decision Variables	
$d_{f_i,f_j}^{m_i,m_j}$	Delay propagated from f_i which is flight of aircraft i which is experienced m_i minute of independent departure delay and f_j is flight of aircraft j with m_j minute of independent departure delay which is possible to swap with aircraft i.
x_{f_i}	Amount of change of scheduled departure time of flight f_i
u_{f_i}	Idle time of flight f_i of aircraft i
s_{f_i,f_j}	Binary variable which equals one if aircraft of flight f_i swap with aircraft of flight f_j, zero otherwise
Parameters	
mtt	Minimum turnaround time of an aircraft
dt_{f_i}	Scheduled departure time of flight f_i of an aircraft i
at_{f_i}	Scheduled arrival time of flight f_i of an aircraft i
$p_{f_i}^{m_i}$	Scheduled arrival time of flight f_i of an aircraft i
$k_{f_i}^{+}$	Maximum limit for the change of departure time moved later
$k_{f_i}^{-}$	Maximum limit for the change of departure time moved earlier
$p_{f_i}^{m}$	Probability of flight f_i with aircraft i experience m minute of independent delay
f_i'	Consecutive flight of flight f_i

The single layer model with swapping option can be stated as:

$$\text{Min} \sum_{i \in A} \sum_{j \in S_{f_i}} \sum_{m_i \in M} \sum_{m_j \in M} \sum_{f_i \in F_i} \sum_{f_j \in F_j} p_{f_i}^{m_i} p_{f_j}^{m_j} d_{f_i,f_j}^{m_i,m_j} \tag{4.14}$$

subject to

$$u_{f_i} = dt_{f_i'} - (1 - \sum_{j \in S_{f_i}} \sum_{f_j \in F_j} s_{f_i,f_j}) at_{f_i} - \sum_{j \in S_{f_i}} \sum_{f_j \in F_j} s_{f_i,f_j} at_{f_j} \quad - x_{f_i} + x_{f_i'} - mtt$$

$$\forall i \in A, \forall f_i \in F_i \tag{4.15}$$

$$d_{f_i,f_j}^{m_i,m_j} \geq 0 \qquad \forall m_i, m_j \in M, \forall f_i \in F_i \quad \forall f_j \in F_j, \forall i \in A, \forall j \in S_{f_i} \tag{4.16}$$

$$-k_{f_i}^{-} \leq x_{f_i} \leq k_{f_i}^{+} \qquad \forall f_i \in F_i, \forall i \in A \tag{4.17}$$

$$u_{f_i} \geq 0 \qquad \forall i \in A, f_i \in F_i \tag{4.18}$$

$$s_{f_i,f_j} = s_{f_j,f_i} \qquad \forall f_i \in F_i, \forall f_j \in F_j \quad \forall i \in A, \forall j \in S_{f_i} \tag{4.19}$$

$$\sum_{j \in S_{f_i}} \sum_{f_j \in F_j} s_{f_j,f_i} \leq 1 \qquad \forall i \in A, \forall f_i \in F_i \tag{4.20}$$

$$d_{f_i,f_j}^{m_i,m_j} = (1 - s_{f_i,f_j})(m_{f_i} - u_{f_i}) + (s_{f_i,f_j})(m_{f_j} - u_{f_i}) \qquad \forall m_i, m_j \in M, \forall f_i \in F_i \quad \forall f_j \in F_j, \forall i \in A, \forall j \in S_{f_i} \tag{4.21}$$

Since swap between two aircraft is possible in the model, the propagated delay should be calculated for each minute for each flight separately. In this model, decision variables u_{f_i}, x_{f_i}, and $d_{f_i,f_j}^{m_i,m_j}$ are integer variables, and s_{f_i,f_j} is a binary variable. The objective function (4.14) is the product of different delay minutes of different aircraft and resulting propagation. In constraint (4.15), slack is redefined since the swapping option exists. If an aircraft i swap with one of the possible aircraft j then the departure time of the next flight of aircraft j is selected to calculate the slack. Constraints from (4.16) to (4.18) are similar to previous models.

(4.19) is the constraint to keep symmetric swapping operation. The constraint (4.20) avoids swapping one flight of an aircraft with another aircraft more than one. (4.21) defines delay propagation according to whether a swap occurs or not. Note that (4.21) is a non-linear equation. If there is a swapping between flight f_i and f_j, delay propagated from the next flight of flight f_i is calculated by using the independent departure delay of f_j, namely m_{f_j}. The single layer model with swapping option is a mixed integer quadratically constrained program since there exists a product of two binary variables in the definition of delay propagation (4.21) which is placed in the constraint set. Nevertheless, modern optimization software packages can solve this problem to optimality for small instances. CPLEX converts the mixed integer quadratically constrained program to a mixed integer problem by linearization. In addition, presolver and aggregator tools in CPLEX, decrease the size of the integer program, simplify the constraints and eliminate redundancy.

4.7. Heuristic Model for the Single Layer Model with Swapping Option

The only difference of the model in Section 4.6 is that heuristic model assumes that even if the swapping occurs at flight f_i between aircraft i and j the propagation of f_i will be calculated based on the delay of an aircraft i, namely m_i. Thus, when calculating the expected propagation of flight f_i the delay probabilities of an aircraft i is used even when swapping occurs with an aircraft j. For this model, decision variable for the delay propagation is redefined as:

- $d_{f_i}^{m_i}$ = Delay propagated from f_i which is flight of aircraft i which experiences m_i minute of independent departure delay and f_j is flight of aircraft j which is possible to swap with aircraft i.

The heuristic model can be stated as:

$$\text{Min} \sum_{i \in A} \sum_{m_i \in M} \sum_{f_i \in F_i} p_{f_i}^{m_i} d_{f_i}^{m_i} \tag{4.22}$$

subject to

$$u_{f_i} = dt_{f'_i} - (1 - \sum_{j \in S_{f_i}} \sum_{f_j \in F_j} s_{f_i,f_j}) at_{f_i} - \sum_{j \in S_{f_i}} \sum_{f_j \in F_j} s_{f_i,f_j} at_{f_j} \quad - x_{f_i} + x_{f'_i} - mtt$$

$$\forall i \in A, \forall f_i \in F_i \tag{4.23}$$

$$d_{f_i}^{m_i} \geq 0 \qquad \forall m_i \in M, \forall f_i \in F_i, \forall i \in A \tag{4.24}$$

$$-k_{f_i}^{-} \leq x_{f_i} \leq k_{f_i}^{+} \qquad \forall f_i \in F_i, \forall i \in A \tag{4.25}$$

$$u_{f_i} \geq 0 \qquad \forall i \in A, f_i \in F_i \tag{4.26}$$

$$s_{f_i,f_j} = s_{f_j,f_i} \qquad \forall f_i \in F_i, \forall f_j \in F_j$$

$$\forall i \in A, \forall j \in S_{f_i} \tag{4.27}$$

$$\sum_{j \in S_{f_i}} \sum_{f_j \in F_j} s_{f_j,f_i} \leq 1 \qquad \forall i \in A, \forall f_i \in F_i \tag{4.28}$$

$$d_{f_i}^{m_i} = m_{f_i} - u_{f_i} \qquad \forall m_i \in M, \forall i \in A, \forall f_i \in F_i \tag{4.29}$$

The difference between the model in Section 4.6 and the model in Section 4.7 is the constraints (4.21), (4.29), and the objective function with a redefined decision variable $d_{f_i}^{m_i}$. Note that the resulting model is a linear MIP. The heuristic model is an approximation of the model in Section 4.6. Since the only constraint that is different than the model in Section 4.6 is the constraint (4.29) (which is the definition of delay propagation when there is no swapping action), there is no infeasible solution for the heuristic model. The reliability of the results of the heuristic model will be discussed in the computational study. Note that, this model is easier to solve computationally.

By using these models, the aim is to avoid the inverse effect of independent departure delay and delay propagation. All these models try to decrease the effect of disruption before it occurs by using predicted independent departure delay probabilities of flights. The objective function is the minimization of total expected delay propagation for the selected aircraft at the single layer model and the multi layer model. For the model with the swapping option, the objective function is the minimization of total expected delay propagation for all aircraft in the schedule by changing the aircraft and crew of a flight.

5. COMPUTATIONAL STUDY

To investigate the improvement in the objective function (expected total delay propagation), three different aircraft are selected from the data. Related information about flights of an aircraft can be seen in Tables 5.1, 5.2, and 5.3. In addition to the schedule information, wind and cloud information are also included since they are used to calculate independent departure delay probabilities of flights. Since scheduled departure and departure time exists in the schedule, the slack of each flight is calculated by assuming the minimum turnaround time is 35 minutes for each flight. For the sake of the computational study, only narrow body planes are selected since there are a lot of idle times with a fewer number of daily flights in the schedule of wide body aircraft. For this reason, there are not many operational solutions to cope with delay propagation for wide body aircraft by changing departure time or swapping aircraft. On the other hand, in general, narrow body planes have relatively tight schedules with more daily flights.

Table 5.1. Daily schedule of TCLSD.

Date	Traffic	Registration	Aircraft	Flight	From	To	STD	STA	Slack	Wind (m/s)	Cloud
6.01.2020	Domestic	TCLSD	Airbus A321neo	TK-2839	TZX	IST	03:45	06:00	35	3	No significant Cloud
6.01.2020	International	TCLSD	Airbus A321neo	TK-794	IST	TLV	07:10	09:25	20	19	Scattered Clouds
6.01.2020	International	TCLSD	Airbus A321neo	TK-795	TLV	IST	10:20	12:50	0	4	No significant Cloud
6.01.2020	International	TCLSD	Airbus A321neo	TK-1505	IST	NUE	13:25	16:30	20	21	Scattered Clouds
6.01.2020	International	TCLSD	Airbus A321neo	TK-1506	NUE	IST	17:25	20:10	70	3	No significant Cloud
6.01.2020	International	TCLSD	Airbus A321neo	TK-368	IST	TAS	21:55	02:40	-	15	Scattered Clouds

Table 5.2. Daily schedule of TCJTA.

Date	Traffic	Registration	Aircraft	Flight	From	To	STD	STA	Slack	Wind (m/s)	Cloud
6.01.2020	International	TCJTA	Airbus A321-200	TK-1258	KBP	IST	01:20	03:25	60	3	Few Clouds
6.01.2020	International	TCJTA	Airbus A321-200	TK-1813	IST	NCE	05:00	08:10	20	15	Scattered Clouds
6.01.2020	International	TCJTA	Airbus A321-200	TK-1814	NCE	IST	09:05	12:00	70	6	Few Clouds
6.01.2020	International	TCJTA	Airbus A321-200	TK-415	IST	VKO	13:45	16:50	20	21	Scattered Clouds
6.01.2020	International	TCJTA	Airbus A321-200	TK-416	VKO	IST	17:45	21:05	60	5	Overcast
6.01.2020	International	TCJTA	Airbus A321-200	TK-792	IST	TLV	22:40	00:50	-	14	Scattered Clouds

Table 5.3. Daily schedule of TCJVJ.

Date	Traffic	Registration	Aircraft	Flight	From	To	STD	STA	Slack	Wind (m/s)	Cloud
6.01.2020	International	TCJVJ	Boeing 737-800	TK-807	ISU	IST	01:00	04:00	160	1	Broken Clouds
6.01.2020	Domestic	TCJVJ	Boeing 737-800	TK-2460	IST	ADA	07:15	09:00	20	19	Scattered Clouds
6.01.2020	Domestic	TCJVJ	Boeing 737-800	TK-2461	ADA	IST	09:55	11:55	25	6	Scattered Clouds
6.01.2020	International	TCJVJ	Boeing 737-800	TK-1775	IST	RIX	12:55	16:00	20	18	Scattered Clouds
6.01.2020	International	TCJVJ	Boeing 737-800	TK-1776	RIX	IST	16:55	19:55	120	3	No significant Cloud
6.01.2020	International	TCJVJ	Boeing 737-800	TK-696	IST	HBE	22:30	00:45	-	15	Scattered Clouds

At first, the probability of experiencing m minute of delay should be calculated for each flight. Note that independent departure delay m has ranged from 0 to 60. There is an upper bound for the delay minutes since the delay above 60 minutes is not absorbed by the scheduled idle time easily. Computation time increases exponentially when the range of m gets bigger. The selected day for calculation, namely 06.01.2020 is a day with extreme weather conditions in Istanbul Airport which is a hub airport of Turkish Airlines. The day with extreme weather condition is selected since it is hypothesized that the schedule slacks will be clustered at İstanbul departure flights after rescheduling. For each flight of the selected three aircraft, related probabilities are calculated by using multivariate kernel density estimation as explained in Section 3.2. The optimization models explained in Chapter 4 are solved by using CPLEX 12.0 solver.

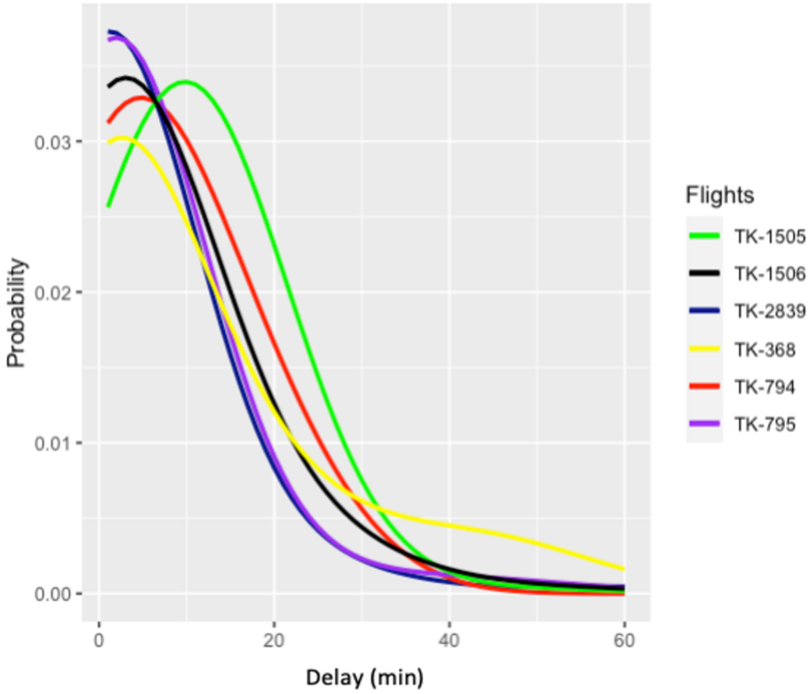


Figure 5.1. Probabilities of daily flights of TCLSD with weighted sample.

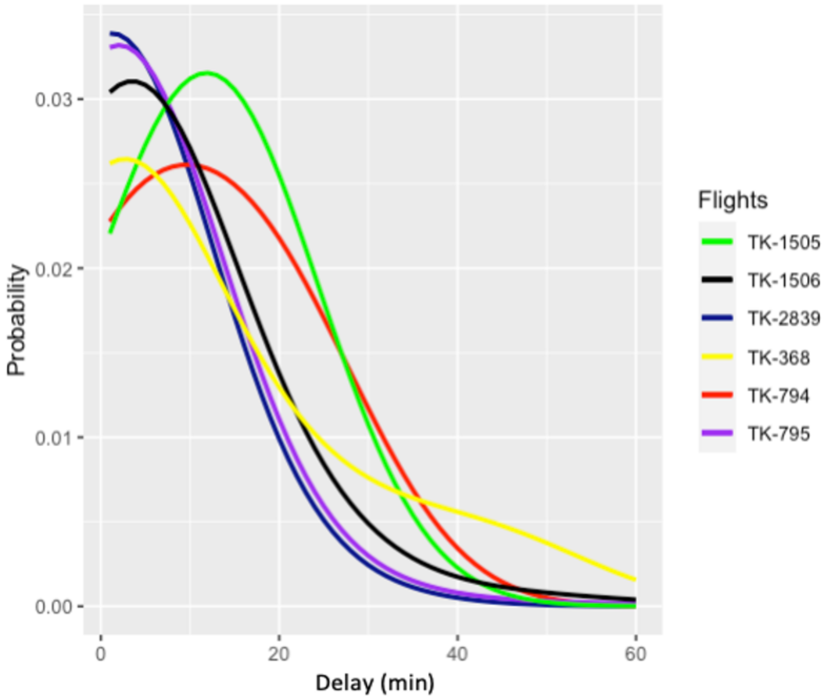


Figure 5.2. Probabilities of daily flights of TCLSD with randomized sample.

5.1. Results for the Single Layer Model

In Figures 5.1 and 5.2, probability plot of flights of TCLSD is presented. Figure 5.1 is created with values calculated from the weighted sample, whereas Figure 5.2 is calculated from the randomized sample space. As expected, flights that have Istanbul as a departure station have higher delay probability values, namely the flights TK-1505, TK-794, TK-368 in both randomized and weighted samples. Since these flights have a higher chance to propagate the delay, slack in the schedule should be transferred to these three Istanbul departure flights.

It can be seen from Table 5.1, aircraft TCLSD has a total of 145 minutes total daily idle time in the originally planned schedule. As explained before the total daily idle time of an aircraft is preserved after flight departure time rescheduling and aircraft swapping operations. Thus there will be a total of 145 minutes total idle time at the end of the optimized model. The third flight of the day for TCLSD, the flight with number TK-795, has 0 slack which means that the flight is extremely vulnerable that even a minute of independent delay will propagate to subsequent flights of TCLSD. At first, the slack of successor and predecessor flights of TK-795 should be delivered to TK-795 to increase robustness. After the slack of the third flight increase until at some point, then the remaining slack of the system is redistributed according to the chance of experiencing departure delay of each flight.

In Tables 5.4 and 5.5 the results for the randomized and weighted sample are presented for the aircraft with registration code TCLSD. Note that, there is no computational burden of a single layer model. In the fourth column, the first number in the square bracket stands for the slack of the first flight, similarly the second number for the second flight, and so on. There are five slack values stated in Table 5.4 in the fourth column since the idle time of the last flight of the day is meaningless. Even 5 minutes allowance of departure time change decreases the total expected delay propagation 45 percent for the weighted sample, 39 percent for the randomized sample. Additionally, increasing the limit on departure time change to 10 and 15 minutes also improves the performance of the single layer model.

On the other hand, even the limit on the departure time change is increased uniformly, namely by 5 minutes, improvement percentage is decreased at higher departure time change limits. This case is caused by, total slack time is kept after rescheduling. Thus it is impossible to totally cut off the total daily delay propagation when the total slack is insufficient.

Table 5.4. Result of Single Layer Model with weighted sample for TCLSD.

Weighted Sample Probabilities	Objective Function Value (minute)	Objective Function Value Improvement	Slack of Flights (minute)
No departure time change (Original Schedule)	9.13	-	[35 20 0 20 70]
Departure time change is allowed 5 min.	5.02	45.02%	[30 20 10 20 65]
Departure time change is allowed 10 min.	3.43	62.41%	[25 21 16 23 60]
Departure time change is allowed 15 min.	2.65	70.92%	[20 24 20 26 55]

Table 5.5. Result of Single Layer Model with randomized sample for TCLSD.

Randomized Sample Probabilities	Objective Function Value (minute)	Objective Function Value Improvement	Slack of Flights (minute)
No departure time change (Original Schedule)	10.64	-	[35 20 0 20 70]
Departure time change is allowed 5 min.	6.46	39.29%	[30 20 10 20 65]
Departure time change is allowed 10 min.	4.36	59.06%	[25 23 14 23 60]
Departure time change is allowed 15 min.	3.06	71.23%	[20 27 17 26 55]

As expected, the model increases the robustness of flight with zero slack, namely TK-795, by sliding departure times of successor and predecessor flights. After the robustness of this highly vulnerable flight increases at some point, the slack times are concentrated at flights TK-794 and TK-1505 since these flights are Istanbul departed flights and their probabilities of experiencing m minutes of independent departure delay are relatively higher as it can be seen from Figures 5.1 and 5.2.

In Tables 5.6 and 5.7, results of the single layer model for the aircraft with registration code TCJTA are stated. Since the total daily slack for an aircraft TCJTA is 230, most of the propagation is absorbed by idle times between the flights. Note that, the slack of the first, the third, and the fifth flight is relatively higher than the second and fourth flights of the day. Thus, the distribution of total slack uniformly between flights solves the issue of propagation. The importance of weather conditions is less for the schedules that have sufficient total slack that is distributed uniformly.

Table 5.6. Result of Single Layer Model with weighted sample for TCJTA.

Weighted Sample Probabilities	Objective Function Value (minute)	Objective Function Value Improvement	Slack of Flights (minute)
No departure time change (Original Schedule)	2.65	-	[60 20 70 20 60]
Departure time change is allowed 5 min.	0.61	77.09%	[55 30 60 30 55]
Departure time change is allowed 10 min.	0.15	94.53%	[50 40 50 40 50]
Departure time change is allowed 15 min.	0.11	96.04%	[46 44 45 46 49]

Table 5.7. Result of Single Layer Model with randomized sample for TCJTA.

Randomized Sample Probabilities	Objective Function Value (minute)	Objective Function Value Improvement	Slack of Flights (minute)
No departure time change (Original Schedule)	3.14	-	[60 20 70 20 60]
Departure time change is allowed 5 min.	0.71	77.29%	[55 30 60 30 55]
Departure time change is allowed 10 min.	0.13	95.83%	[50 40 50 40 50]
Departure time change is allowed 15 min.	0.08	97.52%	[45 47 43 44 51]

As explained, the homogeneous distribution of total daily slack may be more crucial than total daily slack when there exists relatively higher daily schedule slack. Since there is a limit for the departure time change, the slack of a flight can be transferred to its predecessor and successor flights until some threshold. The results of single layer model for the aircraft TCJVJ can be seen in Tables 5.8 and 5.9. For the TCJVJ, the total daily slack is 345 minutes which is much more than the total daily slack of TCJTA. On the other hand, the expected total daily propagation and improvement levels are better for TCJTA than TCJVJ. This is caused by the slack distribution of TCJTA is more homogeneous than TCJVJ.

Table 5.8. Result of Single Layer Model with weighted sample for TCJVJ.

Weighted Sample Probabilities	Objective Function Value (minute)	Objective Function Value Improvement	Slack of Flights (minute)
No departure time change (Original Schedule)	4.98	-	[160 20 25 20 120]
Departure time change is allowed 5 min.	2.66	46.53%	[155 23 22 30 115]
Departure time change is allowed 10 min.	1.56	68.56%	[150 23 22 40 110]
Departure time change is allowed 15 min.	1.06	78.65%	[145 24 24 47 105]

Table 5.9. Result of Single Layer Model with randomized sample for TCJVJ.

Randomized Sample Probabilities	Objective Function Value (minute)	Objective Function Value Improvement	Slack of Flights (minute)
No departure time change (Original Schedule)	4.25	-	[160 20 25 20 120]
Departure time change is allowed 5 min.	1.95	54.03%	[155 21 24 30 115]
Departure time change is allowed 10 min.	0.90	78.95%	[150 21 24 40 110]
Departure time change is allowed 15 min.	0.47	88.97%	[145 22 26 47 105]

5.2. Results for the Multi Layer Model

The results of multi layer model for the aircraft with registration code TCLSD are stated in Tables 5.10 and 5.11. Note that, it is assumed that the delay propagates until it is fully absorbed in the multi layer model. Thus, it is a more realistic model than the single layer model. As expected, the expected total daily delay propagation is higher at the multi layer model. On the other hand, the difference between the results of the multi layer and the single layer model decreases when the limit on the change of departure time increases. The redistribution of slack times for the multi layer model is almost the same as the single layer model. The results of the multi layer model for aircraft TCJTA and TCJVJ can be found in the Appendix.

Table 5.10. Result of Multi Layer Model with weighted sample for TCLSD.

Weighted Sample Probabilities	Objective Function Value (minute)	Objective Function Value Improvement	Slack of Flights (minute)
No departure time change (Original Schedule)	10.98	-	[35 20 0 20 70]
Departure time change is allowed 5 min.	5.59	49.10%	[30 20 10 20 65]
Departure time change is allowed 10 min.	3.67	66.58%	[25 20 17 23 60]
Departure time change is allowed 15 min.	2.78	74.65%	[20 24 20 26 55]

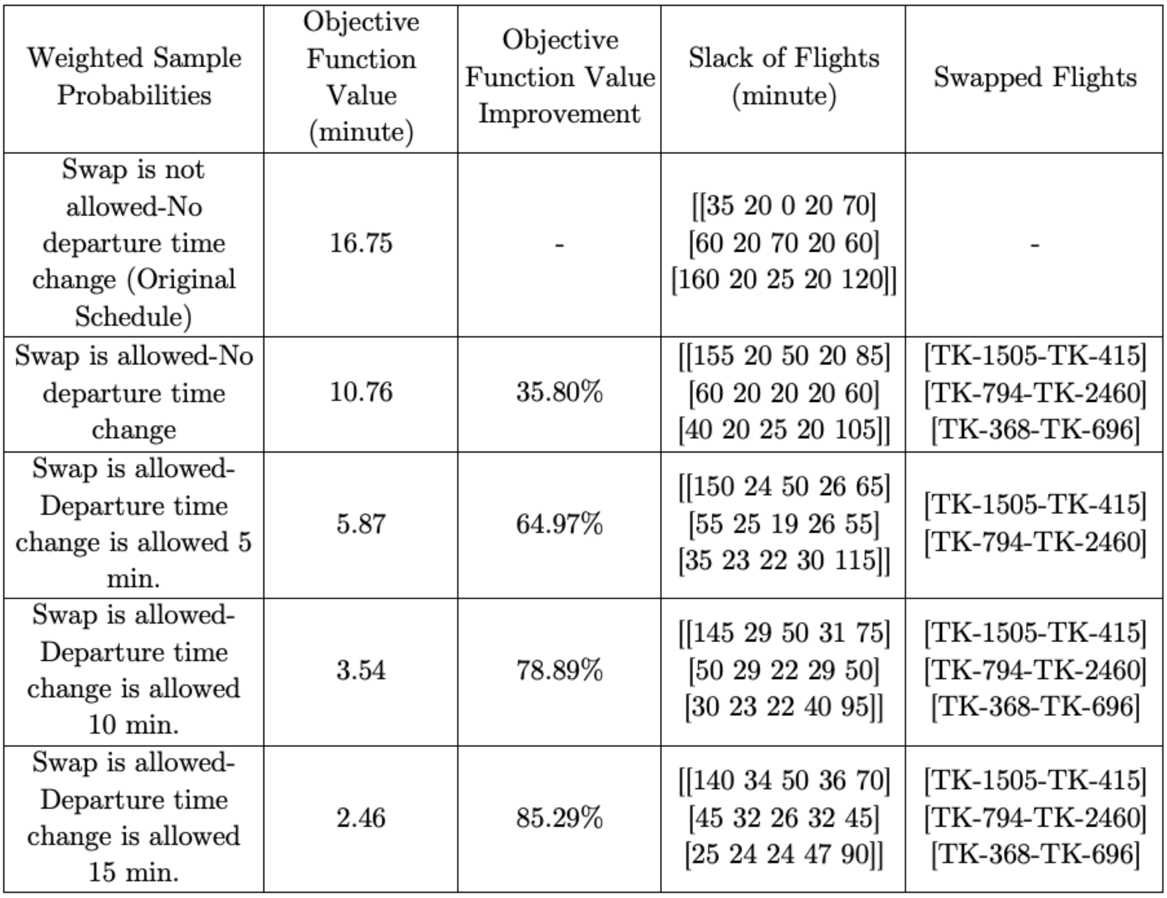
Table 5.11. Result of Multi Layer Model with randomized sample for TCLSD.

Randomized Sample Probabilities	Objective Function Value (minute)	Objective Function Value Improvement	Slack of Flights (minute)
No departure time change (Original Schedule)	13.60	-	[35 20 0 20 70]
Departure time change is allowed 5 min.	7.20	47.05%	[30 20 10 20 65]
Departure time change is allowed 10 min.	4.58	66.35%	[25 24 14 22 60]
Departure time change is allowed 15 min.	3.13	76.97%	[20 27 17 26 55]

5.3. Results for the Single Layer Model with Swapping Option

As explained in Section 4.1, another solution to cope with delay propagation is swapping the duties of aircraft. By doing this swapping operation, slack in the original schedule will be distributed efficiently. The result of the single layer model with the swapping option can be seen in Tables 5.12 and 5.13. Daily flights of the selected three aircraft with registration codes TCLSD, TCJTA, TCJVJ are considered as closed schedule, and duties of these aircraft can be swapped if it is possible. In the fourth column, the slacks of flights are given. In each cell, there are three square brackets. First bracket is for the aircraft TCLSD, the second for TCJTA, and the third for the TCJVJ. The computation time of the single layer model with swapping option for three aircraft is 4 minutes.

Table 5.12. Result of the SLM with swapping option with the weighted sample.



Weighted Sample Probabilities	Objective Function Value (minute)	Objective Function Value Improvement	Slack of Flights (minute)	Swapped Flights
Swap is not allowed-No departure time change (Original Schedule)	16.75	-	[[35 20 0 20 70] [60 20 70 20 60] [160 20 25 20 120]]	-
Swap is allowed-No departure time change	10.76	35.80%	[[155 20 50 20 85] [60 20 20 20 60] [40 20 25 20 105]]	[TK-1505-TK-415] [TK-794-TK-2460] [TK-368-TK-696]
Swap is allowed-Departure time change is allowed 5 min.	5.87	64.97%	[[150 24 50 26 65] [55 25 19 26 55] [35 23 22 30 115]]	[TK-1505-TK-415] [TK-794-TK-2460]
Swap is allowed-Departure time change is allowed 10 min.	3.54	78.89%	[[145 29 50 31 75] [50 29 22 29 50] [30 23 22 40 95]]	[TK-1505-TK-415] [TK-794-TK-2460] [TK-368-TK-696]
Swap is allowed-Departure time change is allowed 15 min.	2.46	85.29%	[[140 34 50 36 70] [45 32 26 32 45] [25 24 24 47 90]]	[TK-1505-TK-415] [TK-794-TK-2460] [TK-368-TK-696]

Table 5.13. Result of the SLM with swapping option with the randomized sample.

Randomized Sample Probabilities	Objective Function Value (minute)	Objective Function Value Improvement	Slack of Flights (minute)	Swapped Flights
Swap is not allowed-No departure time change (Original Schedule)	18.03	-	[[35 20 0 20 70] [60 20 70 20 60] [160 20 25 20 120]]	-
Swap is allowed-No departure time change	12.14	32.63%	[[155 20 50 20 85] [60 20 20 20 60] [40 20 25 20 105]]	[TK-1505-TK-415] [TK-794-TK-2460] [TK-368-TK-696]
Swap is allowed-Departure time change is allowed 5 min.	6.26	65.29%	[[150 26 50 24 65] [55 24 19 27 55] [35 21 24 30 115]]	[TK-1505-TK-415] [TK-794-TK-2460]
Swap is allowed-Departure time change is allowed 10 min.	3.27	81.86%	[[145 31 50 29 75] [50 28 22 30 50] [30 21 24 40 95]]	[TK-1505-TK-415] [TK-794-TK-2460] [TK-368-TK-696]
Swap is allowed-Departure time change is allowed 15 min.	1.87	89.60%	[[140 36 50 34 70] [45 31 26 33 45] [25 22 26 47 90]]	[TK-1505-TK-415] [TK-794-TK-2460] [TK-368-TK-696]

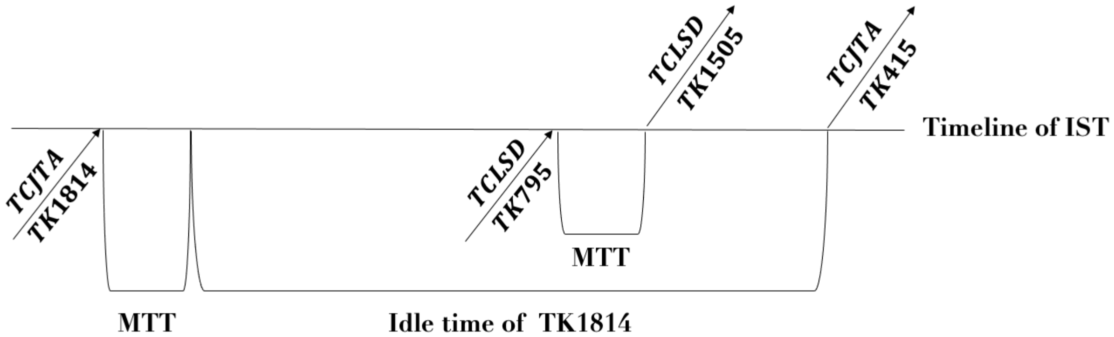


Figure 5.3. Before swapping the planes TK1505 and TK415.

As it can be seen from Table 5.12, even departure time change is set to zero there is a 35.80% improvement in the objective function by swapping the planes. Since the slack of the flight TK795 is 0, the aircraft of flight TK1505, TCLSD, is swapped with the aircraft of flight TK415 namely TCJTA. On the other hand, flight TK1814 has 70 minutes of idle time. Note that, TK1814 and TK415 are consecutive flights of the aircraft with registration code TCJTA. Similarly, TK795 and TK1505 are consecutive flights of TCLSD. After swapping the aircraft of the flight TK1505 and TK415, the slack of TK1814 is transferred to TK795 as in the Figure 5.3 and 5.4. Increasing the allowance of departure time change results in a decrease in the objective function value as expected. By allowing 15 minutes of the departure time change and swapping, the objective function decreases 85.29% for the weighted sample.

The model doesn't have to decide swapping always for the cases like TK1505 and TK415 in Figure 5.3. In Figure 5.3, the aircraft (TCJTA) that arrive to the airport first depart from the airport last. Whereas, the aircraft (TCLSD) that arrive the airport last depart from the airport first. This is an example of the nonhomogenous distribution of idle time between planes. If flight TK1814 experiences unexpected weather conditions at its departure airport, mechanical failures, or any other situation which leads to longer independent delay, then there should be large enough idle times to avoid delay propagation. Also, the conditions for the departure should be ideal for flight TK795 to depart at the planned time since it has zero slack. On the other hand, since predicting the independent delay of these types of extreme and rare cases is difficult, theoretical decisions may not overlap with real life scenarios.

When swapping and 5 minutes of the departure time change are allowed, the flights TK368 and TK696 are not swapped, unlike other levels. At the 5 minutes of departure time change allowance level (third row), even swap does not occur, the slack of two flights, TK368 and TK696, is higher than the defined maximum independent delay length (60 minutes). Thus, there will be no difference for the objective function value either swap occurs or not. However, swapping planes of the flights TK368 and TK696 may be useful for the cases when longer independent delays occur as in the real cases.

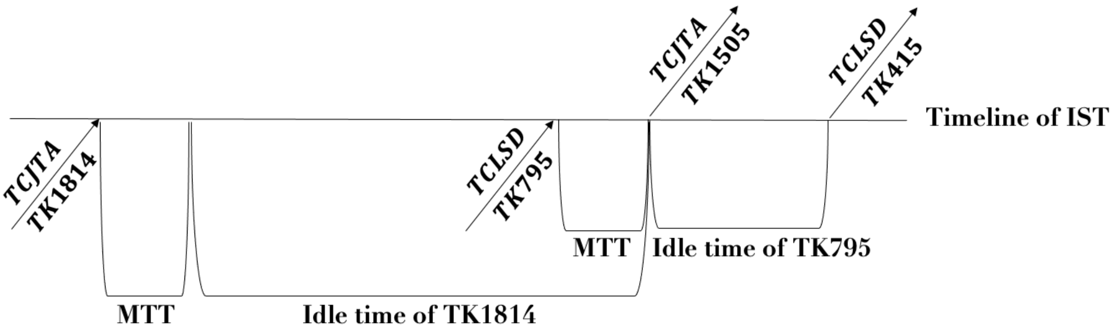


Figure 5.4. After swapping the planes of TK1505 and TK415.

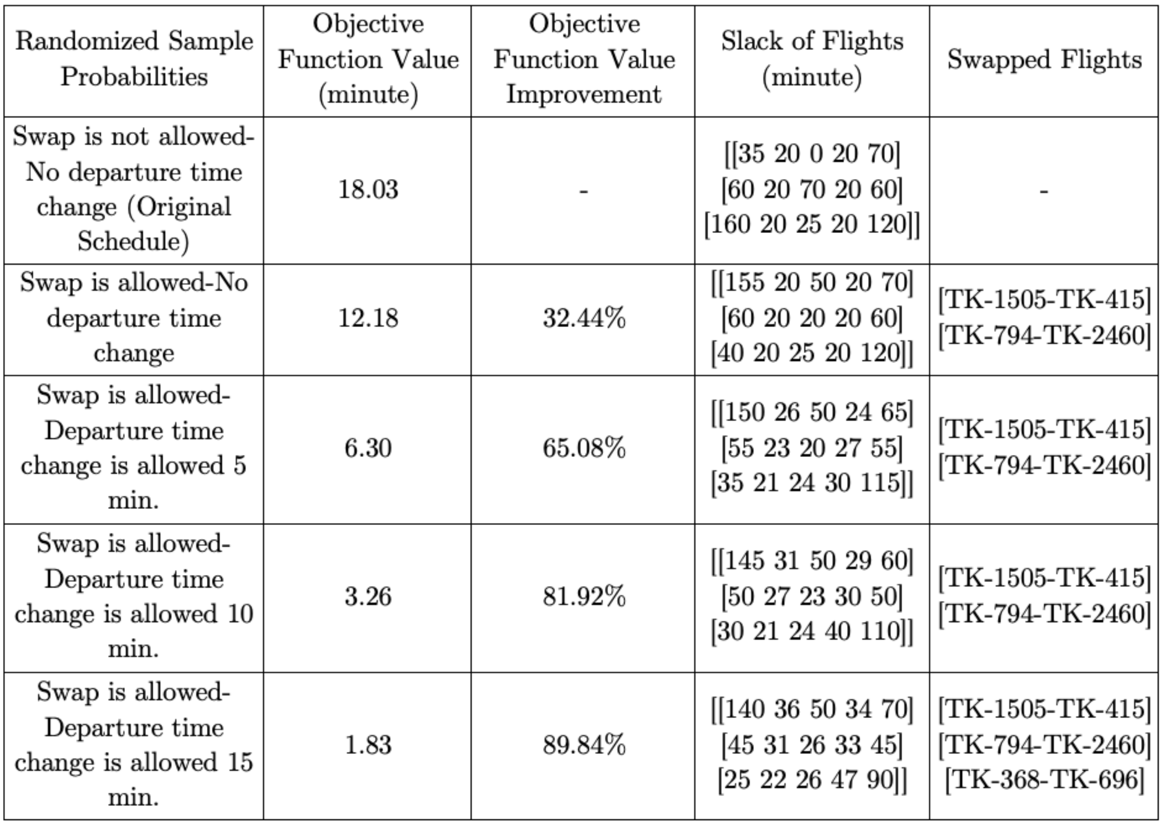
5.4. Results for the Heuristic Model for the Single Layer Model with Swapping Option

In Tables 5.14 and 5.15, the results for the heuristic model for the single layer with swapping options for previously selected three aircraft are presented. The results are close to the result of the exact model as in Tables 5.12 and 5.13. Swapping decisions are the same except that the last flight of the first aircraft and the last flight of the third aircraft for some allowance levels but this does not change the objection function value as explained. Redistributed slack of flights are almost identical with the exact model.

Table 5.14. Result of SLM Heuristic with swapping option with weighted sample.

Weighted Sample Probabilities	Objective Function Value (minute)	Objective Function Value Improvement	Slack of Flights (minute)	Swapped Flights
Swap is not allowed-No departure time change (Original Schedule)	16.75	-	[[35 20 0 20 70] [60 20 70 20 60] [160 20 25 20 120]]	-
Swap is allowed-No departure time change	10.56	36.99%	[[155 20 50 20 70] [60 20 20 20 60] [40 20 25 20 120]]	[TK-1505-TK-415] [TK-794-TK-2460]
Swap is allowed-Departure time change is allowed 5 min.	5.63	66.40%	[[150 24 50 26 65] [55 26 18 26 55] [35 23 22 30 115]]	[TK-1505-TK-415] [TK-794-TK-2460]
Swap is allowed-Departure time change is allowed 10 min.	3.29	80.35%	[[145 29 50 31 60] [50 29 22 29 50] [30 23 22 40 110]]	[TK-1505-TK-415] [TK-794-TK-2460]
Swap is allowed-Departure time change is allowed 15 min.	2.18	86.97%	[[140 34 50 36 70] [45 32 26 32 45] [25 24 24 47 90]]	[TK-1505-TK-415] [TK-794-TK-2460] [TK-368-TK-696]

Table 5.15. Result of SLM Heuristic with swapping option with randomized sample.



Randomized Sample Probabilities	Objective Function Value (minute)	Objective Function Value Improvement	Slack of Flights (minute)	Swapped Flights
Swap is not allowed-No departure time change (Original Schedule)	18.03	-	[[35 20 0 20 70] [60 20 70 20 60] [160 20 25 20 120]]	-
Swap is allowed-No departure time change	12.18	32.44%	[[155 20 50 20 70] [60 20 20 20 60] [40 20 25 20 120]]	[TK-1505-TK-415] [TK-794-TK-2460]
Swap is allowed-Departure time change is allowed 5 min.	6.30	65.08%	[[150 26 50 24 65] [55 23 20 27 55] [35 21 24 30 115]]	[TK-1505-TK-415] [TK-794-TK-2460]
Swap is allowed-Departure time change is allowed 10 min.	3.26	81.92%	[[145 31 50 29 60] [50 27 23 30 50] [30 21 24 40 110]]	[TK-1505-TK-415] [TK-794-TK-2460]
Swap is allowed-Departure time change is allowed 15 min.	1.83	89.84%	[[140 36 50 34 70] [45 31 26 33 45] [25 22 26 47 90]]	[TK-1505-TK-415] [TK-794-TK-2460] [TK-368-TK-696]

The heuristic model which is explained in Section 4.7 gives solutions in a considerably short time which is what planning departments are seeking. On the other hand, in some conditions, it may give wrong swapping decisions. Since the heuristic model calculates delay propagation without considering swap decisions (see the constraint 4.29), when the weather conditions of departure airports (departure airport of the flight leg when the arrival airport is İstanbul) of flights that are likely to swap is significantly different, there may be wrong swap decisions taken by the heuristic model. Thus, if there is an airport with extreme weather conditions, then it may be logical to exclude it from the schedule. Flights that depart from the excluded airport may be rearranged by trusting the experience of the planning department. Then schedule which does not include the airport with extreme weather conditions can be rearranged by the heuristic model which can find near optimal solutions.

5.5. Application of the heuristic model for the whole schedule

Since the heuristic model is a linear MIP, we can solve larger instances of the problem. The heuristic model for the single layer model with swapping options is applied to the whole schedule of Turkish Airlines at 06.01.2020. Note that, there are 51 planes included in the model. There are 152 narrow body aircraft in the original schedule which has at least one daily flight at 06.01.2020. However only planes that have at least five flights in a day are selected. Planes that have four or fewer daily flights have relatively long idle times in their daily flight schedule. That long idle time may be located into the schedule for planned maintenance or any other required duty. Since planned maintenance and duties information is not included in our data, we extract 101 planes which have four or less total daily flight from the data. In addition, crew duty time restrictions may be violated by swapping planes that have four or less daily flights with planes that have five or six daily flights. This is caused by one of the daily flights the planes with four daily flights may be a relatively long flight.

Selected 51 planes fly a total of 282 flights at 06.01.2020. Half of the flights depart from İstanbul Airport. Since the swapping option is defined only at İstanbul Airport, there is a total of 141 flights that can be swapped. Results are listed in Tables 5.16 and 5.17. The total number of swapped flights is between 68 and 76 that depends on the level of departure time change allowance. Approximately half of the flights that depart from İstanbul are swapped by the heuristic model. The number of swapped flights may decrease if crew duty time restrictions are included in the data. Since there is no additional cost of aircraft swapping, the number of swapped flights in Tables 5.16 and 5.17 can be acceptable to decrease the expected total daily delay propagation. Problem is solved in 50 seconds are calculated on Intel® Core® i5 1.6 GHz CPU for the schedule with 51 aircraft. Even the size of the schedule is increased, the improvement of objective function value is similar. When the swap is allowed and departure time change permission is 15 minutes, improvement in total expected delay propagation is around 90%.

Table 5.16. Result of SLM Heuristic with swapping option with weighted sample.

Weighted Sample Probabilities	Objective Function Value (minute)	Objective Function Value Improvement	Number of Swapped Flights
Swap is not allowed-No departure time change (Original Schedule)	211.74	-	-
Swap is allowed-No departure time change	167.60	20.84%	72
Swap is allowed-Departure time change is allowed 5 min.	71.53	66.22%	72
Swap is allowed-Departure time change is allowed 10 min.	35.51	83.23%	68
Swap is allowed-Departure time change is allowed 15 min.	21.21	89.98%	74

Table 5.17. Result of SLM Heuristic with swapping option with randomized sample.

Randomized Sample Probabilities	Objective Function Value (minute)	Objective Function Value Improvement	Number of Swapped Flights
Swap is not allowed-No departure time change (Original Schedule)	223.10	-	-
Swap is allowed-No departure time change	177.57	20.41%	72
Swap is allowed-Departure time change is allowed 5 min.	70.99	68.18%	74
Swap is allowed-Departure time change is allowed 10 min.	31.43	85.91%	72
Swap is allowed-Departure time change is allowed 15 min.	17.14	92.32%	76

6. CONCLUSION

In the competitive aviation industry, airline companies try to improve their profitability, quality of service, and customer satisfaction. Airline schedule planning is a very crucial duty for the planning departments of airline companies. Schedule planning has many different levels, aspects, limits, and regulations. While airline companies make their plans, they have an assumption that the plans work as expected. On the other hand, it is almost impossible for the airline industry since there are plenty of factors that will disrupt the plans. At that point, planning departments try to catch the original schedule as fast as possible and take some precautions that may decrease the effects of disruptions.

In this study, flight departure time retiming and aircraft swapping are used to cope with disruptions. Note that, both two options are proactive ways to avoid delay propagation. Since these precautions are proactive options to avoid inverse effects of propagated delay, the probabilities of flight delays are required to calculate total expected delay propagation. In this thesis realized schedule data from Turkish Airlines are used. Probabilities are calculated by using the three month schedule of Turkish Airlines. Since realized and planned departure times and some factors that may affect the delay chance of a flight are known, the underlying delay mechanism of delay is analyzed. By using multivariate conditional kernel density estimation, probabilities of delay for a flight are calculated according to given cloudiness, wind, departure time, and whether the flight is domestic or international. By using the delay probabilities, the expected delay propagation of the aircraft is calculated by using the single layer at first by assuming the independent delay at a flight affects only its direct connection flight. The decision taken by the single layer model is only changing the departure times of flights to reallocate existing slack in the schedule. Even the departure time change allowance is 5 minutes, the improvement in the objective function value is significant. For the multi layer model, it is assumed that the effects of the independent delay continue until it is fully absorbed. There are also significant improvements in the computational results for multi layer model.

Then, the swapping option is added to the single layer model. By creating a closed schedule with three aircraft, the exact solution is obtained. Since the computation time is relatively longer for the swapping models and the solution should be found as fast as possible, a heuristic model is constructed. The results and decisions of the heuristic model for the previously created schedule with three aircraft are parallel with the exact model. Thus, the heuristic model is also applied to the daily schedule of Turkish Airlines with selected 51 aircraft, and expected improvements are observed.

In this research, the total expected delay propagation is decreased without changing the total cost of idle time. The options used in the model to avoid disruption are not expensive as keeping reserve aircraft or crew. Thus it may be useful to use these models in a planning process to cope with delay propagation, especially to avoid relatively shorter independent delays. The model is also more useful for the companies that have relatively shorter planned idle time in their schedule since when there is enough idle time in the schedule propagation is no longer an issue. For the same reason, the models have better performance for narrow body planes.

For future research, different methods can be used to decrease propagated delay. For instance, the symmetricity constraint on aircraft swapping can be relaxed. In this thesis, if aircraft of flight f_i operates the flight f_j, then aircraft of flight f_j must operates the flight f_i. If this symmetricity constraint is relaxed, when the aircraft of flight f_i operates the flight f_j, then there must be another aircraft, not necessarily aircraft of flight f_j, to operate flight f_i. Variable total daily slack may be another approach to the problem. In this thesis, total daily slack is preserved after rescheduling. However, an improved total daily expected delay may be provided by decreasing the total daily slack.

REFERENCES

1. IATA, “Economic Performance of the Airline Industry”, `https://www.iata.org/en/iata-repository/publications/economic-reports/airline-industry-economic-performance---november-2020---report/`, Accessed on March 2021.

2. ICAO, “The World of Air Transport in 2019”, `https://www.icao.int/annual-report-2019/Pages/the-world-of-air-transport-in-2019.aspx`, Accessed on March 2021.

3. FAA, “Cost of Delay Estimates”, `https://www.faa.gov/data_research/aviation_data_statistics/media/cost_delay_estimates.pdf`, Accessed on March 2021.

4. ICAO, “Air passenger volumes 2019”, `https://www.iata.org/en/iata-repository/publications/economic-reports/Infographic_Economics_Air_Passenger_Demand_2019/`, Accessed on March 2021.

5. Barnhart, C. and K. Talluri, *Airline Operations Research. Design and Operation of Civil and Environmental Engineering Systems*, chap. 10, New York: Wiley, 1997.

6. Marla, L., B. Vaaben and C. Barnhart, “Integrated Disruption Management and Flight Planning to Trade Off Delays and Fuel Burn”, *51st AGIFORS Annual Proceedings - Annual Symposium and Study Group Meeting, AGIFORS 2011*, Vol. 1, 01 2012.

7. “Delay cause by year,percent of total delay minutes”, `https://www.bts.gov/delay-cause-year-percent-total-delay-minutes`, Accessed on March 2021.

8. Barnhart, C., P. Belobaba and A. R. Odoni, “Applications of Operations Research in the Air Transport Industry”, *Transportation Science*, Vol. 37(4), pp. 368–391,

2003.

9. Lan, S., J.-P. Clarke and C. Barnhart, “Planning for Robust Airline Operations: Optimizing Aircraft Routings and Flight Departure Times to Minimize Passenger Disruptions”, *Transportation Science*, Vol. 40:1, pp. 15–28, 2006.

10. Kohl, N., A. Larsen, J. Larsen, A. Ross and S. Tiourine, “Airline disruption management—Perspectives, experiences and outlook”, *Journal of Air Transport Management*, Vol. 13, pp. 149–162, 05 2007.

11. Stojković, G., F. Soumis, J. Desrosiers and M. Solomon, “An optimization model for a real-time flight scheduling problem”, *Transportation Research Part A: Policy and Practice*, Vol. 36, pp. 779–788, 02 2002.

12. AhmadBeygi, S., A. Cohn and M. Lapp, “Decreasing airline delay propagation by re-allocating scheduled slack”, *IIE Transactions*, Vol. 42(7), p. 478–489, 2010.

13. Chiraphadhanakul, V. and C. Barnhart, “Robust flight schedules through slack re-allocation”, *EURO Journal on Transportation and Logistics*, Vol. 2, 11 2013.

14. Aloulou, M., M. Haouari and F. Zeghal Mansour, “Robust Aircraft Routing and Flight Retiming”, *Electronic Notes in Discrete Mathematics*, Vol. 36, pp. 367–374, 08 2010.

15. Deshpande, V. and M. Arıkan, “The Impact of Airline Flight Schedules on Flight Delays”, *Manufacturing Service Operations Management*, Vol. 14, pp. 423–440, 07 2012.

16. Rosenberger, J., E. Johnson and G. Nemhauser, “Rerouting Aircraft for Airline Recovery”, *Transportation Science*, Vol. 37, pp. 408–421, 11 2003.

17. Aktürk, M. S., A. Atamtürk and S. Gürel, “Aircraft Rescheduling with Cruise Speed Control”, *Operations Research*, Vol. 62, pp. 829–845, 2014.

18. Duran, S., S. Gürel and S. Akturk, “Robust Airline Scheduling with Controllable Cruise Times and Chance Constraints”, *IIE Transactions*, 01 2012.
19. Vos, H.-W., B. Santos and T. Omondi, “Aircraft Schedule Recovery Problem – A Dynamic Modeling Framework for Daily Operations”, *Transportation Research Procedia*, Vol. 10, 07 2015.
20. Cao, Y., C. Zhu, Y. Wang and Q. Li, “A Method of Reducing Flight Delay by Exploring Internal Mechanism of Flight Delays”, *Journal of Advanced Transportation*, Vol. 2019, pp. 1–8, 12 2019.
21. Belcastro, L., F. Marozzo, D. Talia and P. Trunfio, “Using Scalable Data Mining for Predicting Flight Delays”, *ACM Transactions on Intelligent Systems and Technology*, Vol. 8, 01 2016.
22. Choi, S., Y. J. Kim, S. Briceno and D. Mavris, “Prediction of weather-induced airline delays based on machine learning algorithms”, pp. 1–6, 09 2016.
23. Ye, B., B. Liu, Y. Tian and L. Wan, “A Methodology for Predicting Aggregate Flight Departure Delays in Airports Based on Supervised Learning”, *Sustainability*, Vol. 12, p. 2749, 04 2020.
24. Hsiao, C.-Y. and M. Hansen, “Econometric Analysis of U.S. Airline Flight Delays with Time-of-Day Effects”, *Transportation Research Record*, Vol. 1951, pp. 104–112, 01 2006.
25. Abdelghany, K., S. Shah, S. Raina and A. Abdelghany, “A model for projecting flight delays during irregular operation conditions”, *Journal of Air Transport Management*, Vol. 10, pp. 385–394, 11 2004.
26. “Weather for 243 countries of the world”, `https://rp5.ru/`, Accessed on March 2021.

27. Silverman, B. W., *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, London, 1986.

28. Shalizi, C., *Advanced data analysis from an elementary point of view*, Cambridge University Press, 2013.

29. Thomas, G. B., M. D. Weir, J. R. Hass and F. R. Giordano, *Thomas' Calculus Early Transcendentals (11th Edition) (Thomas Series)*, Addison-Wesley Longman Publishing Co., Inc., USA, 2005.

30. Hayfield, T. and J. S. Racine, "Nonparametric Econometrics: The np Package", *Journal of Statistical Software*, Vol. 27, No. 5, 2008, `http://www.jstatsoft.org/v27/i05/`.

31. Sepehr, S., "Minimizing airline passenger delay through integrated flight scheduling and aircraft routing", , 10 2005.

32. Hondet, G., L. Delgado and G. Gurtner, "Airline Disruption Management with Aircraft Swapping and Reinforcement Learning", *SID 2018, 8th SESAR Innovation Days*, Salzburg, Austria, Dec. 2018, `https://hal-enac.archives-ouvertes.fr/hal-01944612`.

33. Cook, A., G. Tanner and A. Lawes, "The Hidden Cost of Airline Unpunctuality", *Journal of Transport Economics and Policy*, Vol. 46, 05 2012.

APPENDIX A: APPLICATION

The results for the multi layer model for the aircraft with registration code TCJTA are presented in Table A.1 and Table A.2. The redistributed slack times and objection function values for different allowance levels in Table A.1 and Table A.2 are the same as the the single layer model results for the same aircraft TCJTA (See Table 5.6 and 5.7). Since there is enough slack in the schedule for an aircraft TCJTA, there is no propagation from the first flight to the third flight (similarly, there is no propagation from the second flight to the fourth flight, and so on).

Table A.2. Result of Multi Layer Model with randomized sample for TCJTA.

Randomized Sample Probabilities	Objective Function Value (minute)	Objective Function Value Improvement	Slack of Flights (minute)
No departure time change (Original Schedule)	3.14	-	[60 20 70 20 60]
Departure time change is allowed 5 min.	0.71	77.29%	[55 30 60 30 55]
Departure time change is allowed 10 min.	0.13	95.83%	[50 40 50 40 50]
Departure time change is allowed 15 min.	0.08	97.52%	[45 47 43 44 51]

Table A.1. Result of Multi Layer Model with weighted sample for TCJTA.

Weighted Sample Probabilities	Objective Function Value (minute)	Objective Function Value Improvement	Slack of Flights (minute)
No departure time change (Original Schedule)	2.65	-	[60 20 70 20 60]
Departure time change is allowed 5 min.	0.61	77.09%	[55 30 60 30 55]
Departure time change is allowed 10 min.	0.15	94.53%	[50 40 50 40 50]
Departure time change is allowed 15 min.	0.11	96.04%	[46 44 45 46 49]

Table A.3. Result of Multi Layer Model with weighted sample for TCJVJ.

Weighted Sample Probabilities	Objective Function Value (minute)	Objective Function Value Improvement	Slack of Flights (minute)
No departure time change (Original Schedule)	5.11	-	[160 20 25 20 120]
Departure time change is allowed 5 min.	2.72	46.74%	[155 22 23 30 115]
Departure time change is allowed 10 min.	1.59	68.81%	[150 23 22 40 110]
Departure time change is allowed 15 min.	1.08	78.89%	[145 24 24 47 105]

Table A.4. Result of Multi Layer Model with randomized sample for TCJVJ.

Randomized Sample Probabilities	Objective Function Value (minute)	Objective Function Value Improvement	Slack of Flights (minute)
No departure time change (Original Schedule)	4.26	-	[160 20 25 20 120]
Departure time change is allowed 5 min.	1.96	54.05%	[155 21 24 30 115]
Departure time change is allowed 10 min.	0.90	78.97%	[150 21 24 40 110]
Departure time change is allowed 15 min.	0.47	88.98%	145 22 26 47 105]

Table A.3 and A.4 are the results of the multi layer model with weighted and random samples for the aircraft TCJVJ. The results for the multi layer model are very close to the results for the single layer model for the aircraft TCJVJ (See Table 5.8 and 5.9).

Table A.5. The first fifteen rows of the data.

Date	Traffic	Registration	Aircraft Age	Aircraft Type	Flight	From	To	Scheduled Departure	Actual Departure	OffBlock	TakeOff	Scheduled Arrival	Actual Arrival	Touch Down	On Block	Departure Delay (minutes)	Arrival Delay (minutes)	Slack (minutes)	Temperature (Celcius)	Wind Direction	Wind Speed(m/s)	Cloud Types	Horisontal Visibility(km)	Dewpoint (Celcius)
5.11.2019	Domestic	TCJRG	12.3	Airbus A321-200	TK-2552	IST	DLM	03:25	03:19	03:19	03:51	05:00	04:54	04:51	04:54	0	0	15	18.00	Wind blowing from the south-west	8.00	No Significant Clouds	10.0 and more	15.00
5.11.2019	Domestic	TCLSJ	0.3	Airbus A321neo	TK-2122	IST	ESB	04:00	03:54	03:54	04:15	05:20	05:04	04:55	05:04	0	0	15	20.00	Wind blowing from the south-west	10.00	No Significant Clouds	10.0 and more	14.00
5.11.2019	Domestic	TCJSI	6.8	Airbus A321-200	TK-2246	IST	GNY	04:30	06:02	06:02	06:28	06:35	07:55	07:52	07:55	92	80	15	20.00	Wind blowing from the south-west	10.00	No Significant Clouds	10.0 and more	14.00
5.11.2019	International	TCJIS	11.4	Airbus A330-200	TK-1523	IST	DUS	04:55	04:55	04:55	05:10	08:25	08:14	08:05	08:14	0	0	50	18.00	Wind blowing from the south-west	7.00	Few Clouds	10.0 and more	15.00
5.11.2019	International	TCLSG	0.8	Airbus A321neo	TK-799	BSR	IST	05:00	04:50	04:50	05:01	08:55	08:23	08:14	08:23	0	0	15	14.00	Wind blowing from the north-east	5.00	Broken Clouds	8.0	11.00
5.11.2019	International	TCJYM	5	Boeing 737-900ER	TK-1981	IST	LGW	05:00	05:00	05:00	05:32	09:20	09:21	09:16	09:21	0	1	35	20.00	Wind blowing from the south-west	10.00	No Significant Clouds	10.0 and more	14.00
5.11.2019	Domestic	TCJTP	3.1	Airbus A321-200	TK-2117	ESB	IST	05:15	05:14	05:14	05:24	06:40	06:30	06:18	06:30	0	0	25	5.00	Calm, no wind	0.00	Few Clouds	10.0 and more	2.00
5.11.2019	International	TCJRB	13.5	Airbus A321-200	TK-1061	IST	LJU	05:20	05:23	05:23	05:52	07:45	07:58	07:53	07:58	3	13	15	18.00	Wind blowing from the south-west	6.00	No Significant Clouds	10.0 and more	15.00
5.11.2019	International	TCJVI	4.1	Boeing 737-800	TK-1881	IST	SKG	05:20	05:24	05:24	05:53	06:55	06:50	06:46	06:50	4	0	15	18.00	Wind blowing from the south-west	6.00	No Significant Clouds	10.0 and more	15.00
5.11.2019	Domestic	TCJRK	11.8	Airbus A321-200	TK-2807	SZF	IST	05:35	05:25	05:25	05:36	07:25	06:50	06:42	06:50	0	0	20	15.00	Wind blowing from the south-west	7.00	No Significant Clouds	10.0 and more	8.00
5.11.2019	Domestic	TCJSU	4.6	Airbus A321-200	TK-2455	ADA	IST	05:45	05:40	05:40	05:59	07:40	07:25	07:16	07:25	0	0	25	10.00	Wind blowing from the north-east	3.00	Scattered Clouds	9.0	13.00
5.11.2019	Domestic	TCJGS	13.6	Boeing 737-800	TK-2217	ADF	IST	05:50	05:46	05:46	05:57	08:00	07:46	07:35	07:46	0	0	20	15.00	Wind blowing from the south-west	9.00	Broken Clouds	10.0 and more	11.00
5.11.2019	International	TCJRM	8.9	Airbus A321-200	TK-1883	IST	VIE	05:50	06:16	06:16	06:38	08:15	08:50	08:40	08:50	26	35	25	18.00	Wind blowing from the south-west	5.00	No Significant Clouds	10.0 and more	15.00
5.11.2019	Domestic	TCJRG	12.3	Airbus A321-200	TK-2553	DLM	IST	05:55	05:48	05:48	06:01	07:25	07:09	07:02	07:09	0	0	85	18.00	Wind blowing from the north	4.00	No Significant Clouds	9.0	9.00