ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/330989073

Aircraft Departure Delay and its Associated Costs for Airlines and Airports in Europe

Research · June 2017

DOI: 10.13140/RG.2.2.34307.55847

CITATIONS
0

READS
1,242

8 authors, including:

Jose Ignacio de Alvear Cardenas
Delft University of Technology
2 PUBLICATIONS **0** CITATIONS

SEE PROFILE

Tobias Haegens
Delft University of Technology
1 PUBLICATION **0** CITATIONS

SEE PROFILE

T. Hendrix
Delft University of Technology
1 PUBLICATION **0** CITATIONS

SEE PROFILE

Fabio Kerstens
Delft University of Technology
1 PUBLICATION **0** CITATIONS

SEE PROFILE

All content following this page was uploaded by Jose Ignacio de Alvear Cardenas on 09 February 2019.

The user has requested enhancement of the downloaded file.

Aircraft Departure Delay and its Associated Costs for Airlines and Airports in Europe

J.I. de Alvear Cárdenas *, T. Haegens *, T. Hendrix *, F. Kerstens *, L. Kokee *, T. van der Duim *, S. van Elsloo *, N. Wildemans *
Delft University of Technology, Delft, Netherlands

Delays, flights that depart at least 5 minutes behind schedule, are a large problem in the aviation industry in terms of cost and management. Significant research has already been conducted in this area in the United States, but much less so in Europe. It is therefore interesting to study delays in Europe. More specifically, the relation between costs and high morning delay have been considered particularly interesting and have thus been investigated. To do so, data from Eurocontrol was obtained which was then filtered to yield a more reliable data set. To study morning delay, a performance parameter for morning delay (PMD) was introduced, which combined the relative amount of delay occurring in the morning with the severity of the delay. Statistical analysis combined with clustering made it possible to determine that airports with more runways generally experience less severe morning delay than others, although there is no apparent relation between airport size and morning delay. Furthermore, it has been found that airports on the western European mainland suffer relatively less from morning delay as opposed to central European and British airports. Additionally, British airports showed to be costly in terms of delay, with London Gatwick being the costliest in all of Europe, whereas central European airports perform better in this respect. A statistically significant relation was found between the PMD and the average cost per passenger for an airport. For further research, it is recommended to use a more advanced data set which includes information regarding arrival delay and tail number.

I. Introduction

The aviation industry is growing rapidly [1], and so is the demand for an increased airspace and airports capacity [2]. To prepare for this increasing demand, the airport capacity needs to be increased[a]. Investigating the occurring delays - where a flight is considered delayed if it departs 5 minutes later than scheduled [3] - can help develop a better understanding of the limitation that delays pose on the capacity. From this capacity, the financial side should not be decoupled. Also in this manner flight delays have a big influence on the capacity, as delays impose a big rise in airline costs limiting the airlines to expand further. Looking at the estimated annual European delay cost of 6.6 to 11.5 billion euros [4], the previous statement could be underlined.

Like any other schedule based transportation system, the European air transport system is sensitive to delays, especially during high season and holidays. Once an initial delay occurs, it is difficult to maintain the schedule and subsequent flights on the same and other airports may be delayed as well; the delay propagates. Many studies have been done in recent years in an effort to study its behaviour and minimize it. Jetzki [5] developed aircraft sequences to study mitigating factors of reactionary delays in Europe. It was shown that reactionary delays at major hub airports primarily affect only their own operations. Fleurquin, Campanelli and Eguiluz [6] developed a more elaborate model based on data from the United States capable of evaluating the risk of system wide congestion. Furthermore, Cook *et al.*[7, 8] conducted research regarding the cost of delay. A numerical relationship between the aircraft delay and the associated cost was established, where the cost per minute delay per passenger increases non-linearly as the delay lengthens.

It has to be noted that the vast majority of delay research is related to the airspace of the United States, where data is obtained from the Bureau of Transportation. Its European counterpart, Eurocontrol, has less detailed, harder obtainable data. Due to the combination of these two factors, most of the research that has been conducted in the United States has not been conducted yet in Europe, which leaves a large research gap in this area. Filling in this gap would lead to a better optimized, more cost-efficient airspace with less delays.

Initially, the objective of this research campaign was to identify flight delay propagation patterns in the European airspace system and to quantify the associated cost. Based on the limitations of the available data we decided instead to evaluate whether European airports that suffer relatively more from severe delay in the morning are also more costly in terms of delay. To show this, two questions should be answered:

*Undergraduate student, Faculty of Aerospace Engineering

[a]Currently most EU airports allow only 80-85% of their capacity to anticipate for possible delays.

- Which airports experience relatively severe delay in the morning and can this be related to airport characteristics?
- Which airlines and airports are financially the most costly in terms of delays and is there a relation between them?

For the first research question, airports with high morning delays will be investigated, as these may well be the root cause of instances of delay propagation. Based on the expected delay costs described, the second research question focuses on distinguishing the cost-efficient and cost-inefficient airports and airlines in Europe. This data can be used to lower future costs.

II. Data description and filtering

Raw data of the flights in European airspace was obtained from Eurocontrol [9]. Specifically, information was provided regarding the departure airport, arrival airport, aircraft type, airline, flight number, requested flight level, and scheduled and actual departure time. The data used in this study was collected arbitrarily for the first day of the months May to December of 2016.

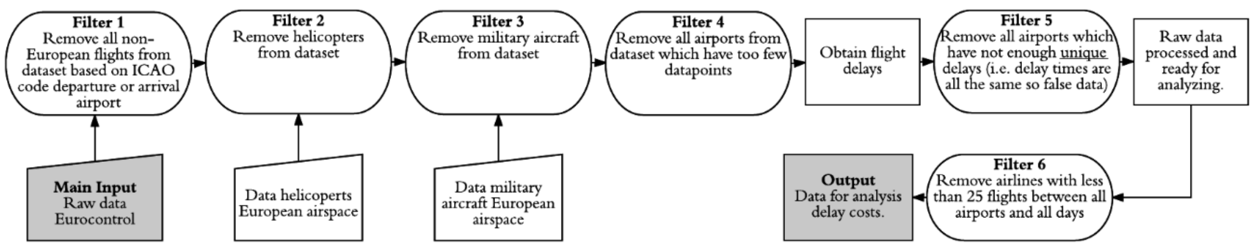

Figure 1. Flowchart of data filtering process.

To answer the research questions, raw data was first processed, using a series of six filters shown in Fig. 1. Filter 1 consisted of removing all flights that did not depart from or arrive at a Eurocontrol member state [10]. The data set also contained flight data on rotorcraft and military aircraft. As these show different cost patterns compared to commercial airlines, they were discarded using filter 2 and filter 3 respectively.

The following two filters were focused on the airports. First, filter 4 consisted of eliminating airports that had too few (< 10) aircraft movements per day as it was considered unlikely statistically significant conclusions could be drawn from them. Then, filter 5 involved eliminating airports which clearly displayed unreliable data. The original data set contained numerous airports which only showed a very small number of different amount of delays. For instance, for some airports, almost every flight on a single delay was either delayed 14 or 15 minutes. Due to these singularities, the data for this airport was considered to be inaccurate. For this reason, for each remaining airport the frequencies of the unique departure delay times were collected. If the number of flights that have the three most common delays (for example: 10, 12 and 15 minutes) made up more than 99% all flights, the airport was discarded from the data set. This threshold was determined using a heuristic approach, where the aim was to filter out all airports where it could be verified using external sources that their data was (partially) incorrect.

Finally, filter 6 was applied exclusively for the second research question. Airlines with fewer than 25 flights over the eight days combined were discarded, as these would be prone to outliers. The threshold of 25 flights was obtained using a heuristic approach.

Table 1. List of airports remaining after the filtering process.

ICAO	Airport	ICAO	Airport	ICAO	Airport
EBBR	Brussels Airport	EGKK	London Gatwick	LIML	Milan Linate
EDDB	Berlin-Schönefeld	EGLC	London City	LIPZ	Venice Marco Polo
EDDF	Frankfurt am Main	EGLL	London Heathrow	LIRF	Rome Fiumicino
EDDL	Düsseldorf	EIDW	Dublin	LKPR	Prague Václav Havel
EDDM	München Franz Josef Stauß	ENGM	Oslo Gardermoen	LSGG	Geneva
EDDS	Stuttgart	LEBL	Barcelona-el Prat	LSZH	Zürich
EFHK	Helsinki-Vantaa	LEMD	Adolfo Suárez Madrid-Barajas	UKBB	Kiev Boryspil
EGCC	Manchester	LFPG	Paris-Charles de Gaulle		
EGGW	London Luton	LIMC	Milan Malpensa		

III. Methodology

A. Morning Delay Performance

The purpose of this subsection is to discuss the methodology employed to investigate which airports experience relatively more morning delay, and which airports more afternoon delays. First, the delay (integer) of all flights was obtained as described in Sec. II. After that, the discrete cumulative distribution function (CDF) of the total delay per airport (that is, the sum of the delay time of each aircraft departing from that airport) could be constructed.

An example of the CDF is shown in Fig. 2, where the fraction of the total departure delay over the day is plotted on the vertical axis[b]. The function reaches a value of unity at the end of the day as all flights and therefore all delays have occurred. The choice for using the CDF comes from the desire to normalize results, so that different airport sizes could be compared.

The moment airport operations start the function increases as delays occur. A quick growth in the beginning of the day indicates high delay in the morning relative to the rest of the day. To give an indication of the growth rate of the CDF, the time difference between the start of operations[c] and the moment when the CDF reaches a 20% threshold was determined for all airports. This time interval will be referred to as ΔT henceforth. In Fig. 2 ΔT is visualized for London Heathrow Airport.

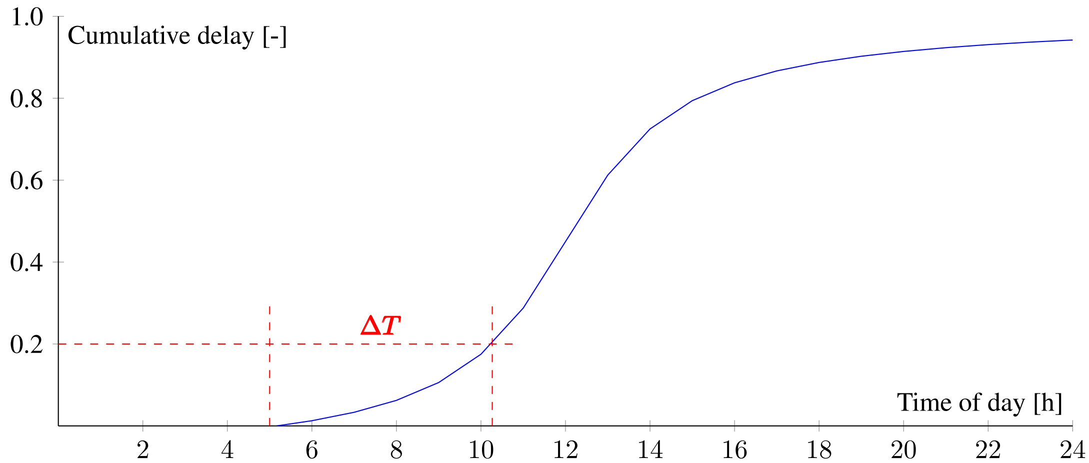

Figure 2. Example of a cumulative distribution function for the total delay. ΔT is indicated by the dashed red lines.

The threshold for ΔT was chosen such that the time window is large enough to obtain significant results, while the time interval still ends in the beginning of the day. Futhermore, analysis performed on varying this threshold from 10% to 30% has not shown big variations in for example airport rankings based on morning delay. In Sec. IV-A the results for other thresholds will be shortly discussed.

So far, the analysis only gives an indication of how early a fraction of the daily delay occurred. As explained by Cook [8], a few flights delayed with considerable delay are more costly and can be considered worse than many flights with low delay. Therefore, the average delay per flight within the time frame of ΔT hours was also computed.

In order to draw conclusions on an airport by airport basis, the data of all eight days was combined, and the mean of ΔT and the average delay per flight were calculated. For a given day, the averaged values of ΔT can be plotted against the averaged delay per flight for all airports under consideration. Plots such as these are shown in Sec. IV-A. The performance of the airport can be considered high when ΔT is large and the average delay per flight is low. A performance parameter for morning delay (PMD) was constructed as shown in Eq. (2), using the ratio of the average delay per flight and the value for ΔT as can be seen in Eq. (1). This way, airports could be ranked easily. A PMD of 0 corresponds to the best performing airport a PMD of 1 to the worst performing airport. In Sec. IV-C it will be evaluated whether higher morning delay has a relationship with the cost on an airport by comparing the PMD parameter with the average cost per passenger due to delay.

[b]For example, if the cumulative delay reaches a value of 0.4, then 40% of the total delay time over that day has been reached.

[c]The moment operations start is taken as the latest moment during the day after which all consecutive flights depart less than one hour apart.

$$x = \frac{\text{average delay per flight}}{\Delta T} \quad (1)$$

$$\text{PMD} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (2)$$

Table 2. Overview of airports grouped by FAA classified size, where annual indicates the annual passenger boardings worldwide in 2016 [11].

Size	Passenger boardings	Airports					
Large Hub	$\geq 1\%$ annual	EGLL	LFPG	EDDF	LEMD	LEBL	EGKK
		EDDM	LIRF				
Medium Hub	$0.25\% \leq$ annual $< 1\%$	EIDW	ENGM	LSZH	EGCC	EDDL	EBBR
		LIMC	EFHK	LSGG	EGGW	LKPR	EDDB
Small Hub	$0.05\% \leq$ annual $< 0.25\%$	EDDS	LIML	LIPZ	UKBB	EGLC	

To see if there was any relation between the airport parameters and morning delay, K-means clustering was applied using SPSS[d]. K-means is an iterative partitional clustering algorithm in which the centroid (cluster center) of each of the groups needs to be specified initially. Between the cluster centers, K bisections are constructed; shaping the clusters' boundaries and dividing the domain in multiple regions [12]. For each of the data points, it is verified in which section they are located such that they are assigned to the cluster with the closest centroid. Once the computations have been carried out for the whole data set, the position of the centroids is recalculated, bearing in mind that they are situated in the middle of all the points that define its cluster. Due to the new clusters' centers, the process is repeated and the points are assigned again to the cluster with the closest centroid. The iteration is carried out until convergence is reached.

Another objective of this analysis is to evaluate what airport parameters have an effect on morning delay performance. To determine whether the relationship were present and statistically significant, Fisher's exact test was used [13]. Two parameters were used; annual passenger flow, as shown in Table 2, and number of runways.

B. Classification of Costs for Airlines

Costs for airlines can be subdivided into soft costs and hard costs. Hard costs are the costs that directly follow from the delay (e.g. passenger compensation if required by government regulations). Soft costs are costs that do not generate an immediate repercussion on the airline; however, they are closely related with the consumers' experience and their future subjective decisions (e.g. passenger does not want to fly with a certain airline any more due to the delays experienced, meaning the airline loses future profit). When answering the research question, only hard costs were considered for the airline cost. This is mostly due to the subjective nature associated with the soft cost. For more information regarding the difference between the two types of costs, the reader is referred to the literature report [14].

In Cook *et al.* [8], the cost per passenger per minute of delay was defined using a non-continuous function, assigning costs only to intervals of delay. To generate a smoother cost function, this function was interpolated using linear splines as suggested by Cook *et al.*, where the middle points of each of these delay intervals were taken as interpolation nodes. The result is tabulated in Table 3 and drawn Fig. 3. As flights that depart less than five minutes later than scheduled are not considered delays, the cost of these flights are set to zero.

The cost per passenger for a given delay time was then easily computed using this function. Then, the cost of operating in a given airport could be estimated for each airline, using a weighted average of the cost per passenger of all flights of that airline that departed from that airport, using the number of passengers of the flights as weights. To determine the number of passengers of each flight, the seating capacity of the aircraft was multiplied with the load factor[e], to take into account that not all aircraft fly at full capacity.

This resulted in the cost per passenger for a given airline at a given airport. It was chosen to use the cost per passenger for *all* passengers, whether they were delayed or not, as this would result the average cost every passenger would pay as part of his ticket price to compensate for the delay cost.

Additionally, the average delay per flight on each day and the standard deviation of the average delay over all eight days were calculated.

[d]Statistical Package for the Social Sciences, used for making statistical models.

[e]This was taken equal to 76.2% for all flights, with 76.2% being the average load factor in European flights [15].

Table 3. Hard cost per minute of delay per passenger for the limits of the delay interval. The load factor was not taken into account.

Delay interval [min]	Cost lower bound [€/min/pass]	Cost upper bound [€/min/pass]
$0 \leq t < 5$	0.0000	0.0000
$5 \leq t < 7.5$	0.0604	0.0907
$7.5 \leq t < 22.5$	0.0907	0.2153
$22.5 \leq t < 37.5$	0.2153	0.2947
$37.5 \leq t < 52.5$	0.2947	0.3513
$52.5 \leq t < 67.5$	0.3513	0.3967
$67.5 \leq t < 82.5$	0.3967	0.4307
$82.5 \leq t < 105$	0.4307	0.5967
$105 \leq t < 150$	0.5967	0.9062
$150 \leq t < 210$	0.9062	1.1668
$210 \leq t < 270$	1.1668	1.2348

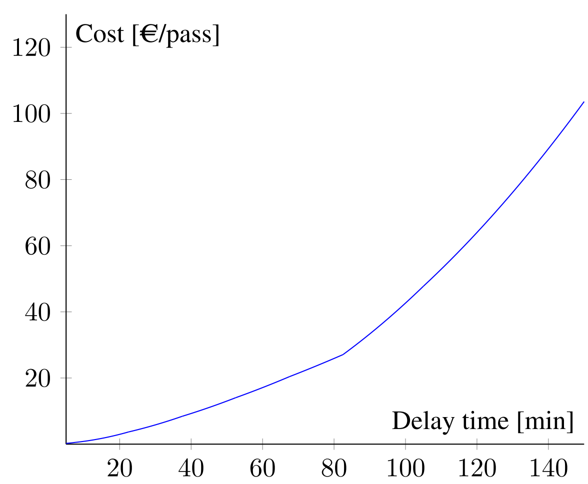

Figure 3. Piece-wise hard cost function per passenger vs minute delay.

IV. Results

A. Morning Delay Performance

Table 4 and Table 5 show the five best and worst performing airports, respectively, according to the PMD number that has been explained in Sec. III-A. The value of PMD indicates the frequency and severity of delayed flights in the morning, a value of unity corresponding to the worst airport.

Table 4. The 5 best performing airports ranked by PMD.

Rank	ICAO	Airport	PMD	ΔT [hr]	Average delay time [min]
1	LKPR	Prague Vaclav Havel	0.0000	5.47	8.49
2	EBBD	Berlin	0.0813	4.04	8.60
3	LIML	Milan Malpensa	0.1117	4.75	11.13
4	UKBB	Kiev Boryspil	0.1199	5.11	12.27
5	LIPZ	Venice Marco Polo	0.1328	4.08	10.16

Table 5. The 5 worst performing airports ranked by PMD.

Rank	ICAO	Airport	PMD	ΔT [hr]	Average delay time [min]
25	EIDW	Dublin	1.0000	2.25	19.35
24	EGKK	London Gatwick	0.8000	2.71	19.55
23	EGCC	Manchester	0.7999	2.57	18.52
22	EGLL	London Heathrow	0.6824	3.08	19.59
21	LSZH	Zurich	0.4717	2.58	12.62

From Table 4 and Table 5, it can be observed the four worst performing airports are located on the British isles, with two being located in London. This might indicate that the delay here is not solely caused by the issues on the airport itself, but the delays may also be caused by a structural problem above London airspace. The fact that London has six airports has a serious impact on the movement of all aircraft flying above this area. However, further investigation will have to be conducted to confirm these hypotheses.

In Fig. 4, the mean time to reach 20% of the daily delay, ΔT, is visualized for each airport on a map of Europe. The airports are indicated by dots, where the size of the dot corresponds to the relative passenger flow per year (a bigger dot means a larger passenger flow). The color corresponds to the value of ΔT.

Several things can be noted from Fig. 4. The airports located in England and Ireland, including the three London airports take a relatively short time to reach the 20% threshold. These airports thus experience relatively high morning delay as reflected in the PMD scores in Table 5. The same phenomenon can be seen for most airports located in central Europe, with the exception of Vaclav Havel Airport Prague (LKPR). On the other hand, the south-western European airports, the ones from Belgium to Spain, have a very low occurrence of morning delay, as well as Boryspil International Airport in Ukraine.

A very obtrusive outlier is London Luton airport, which performs very well despite being close to 3 relatively bad performing airports. This could be due to the combination of it being a relatively small airport and it being located far from the the center of London. Capacity problems are a known cause of delays in Heathrow and Gatwick airport, and London City, being located in the heart of London, causes this airport to experience a lot of delay.

Figures 5 and 6 show the time to reach the threshold of 20% on the horizontal axis. On the vertical axis, the mean delay per flight is shown. In Fig. 5, airports have been grouped on size, (Table 2); in Fig. 6 they have been grouped based on the number of runways available at the airport. The clusters calculated using SPSS are represented by division using the dashed lines.

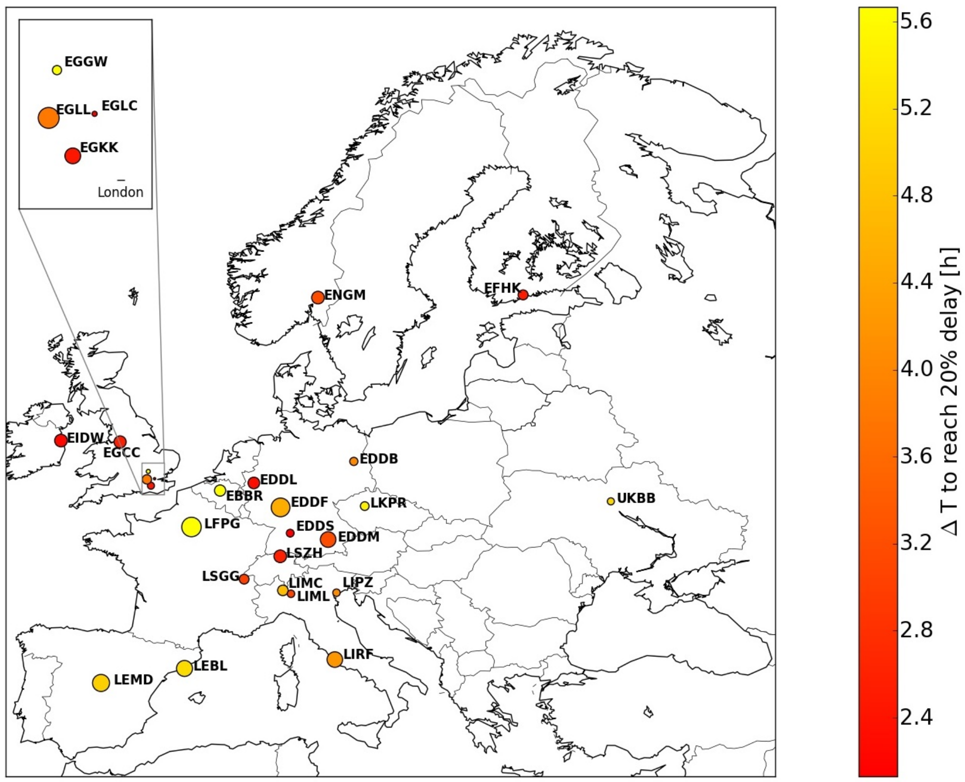

Figure 4. Representation of the time it takes for an airport to reach 20% of its daily delay.

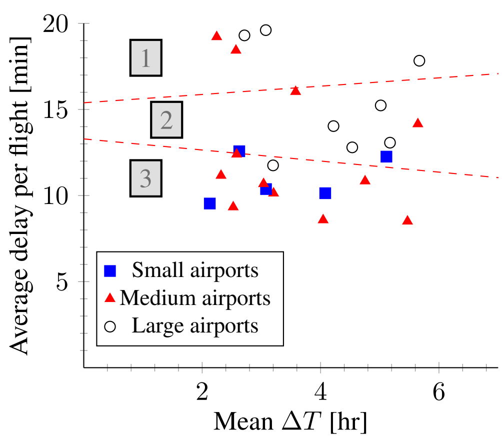

Figure 5. Airports grouped on size of airport.

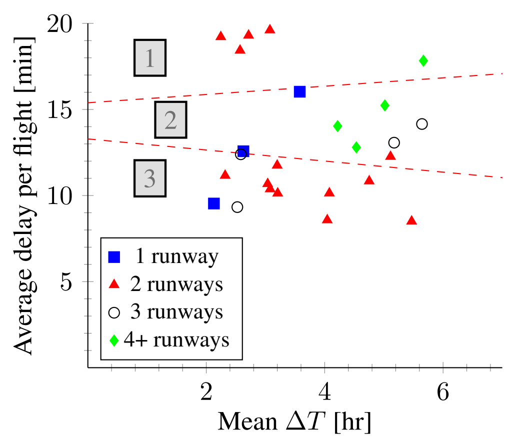

Figure 6. Airports grouped on size of airport.

The airports within cluster 1 all have high average delays per flight in the morning; the airports within cluster 2 have moderate average delay per flight and the airports within cluster 3 have relatively low average delay per flight. However, there is little relation between the clusters and the value of ΔT for airports in the cluster.

Concerning the clustering in Fig. 5, it is difficult to distinguish patterns. A weak pattern can be distinguished in the sense that large airports seem to be concentrated in cluster 2, which would imply that large airports generally have less delay per flight in the morning. However, using Fisher's exact test, this pattern is not statistically significant with $p > 0.05$.

By inspection, in Fig. 6 it can be seen that there is a stronger relationship between the number of runways of an airport and the cluster it belongs to. Again, Fisher's exact test was performed in order to determine if the relationship was statistically significant. The test resulted in a p-value of 0.027, indicating that the relationship is significant. 75% of the airports with more than three runways are a member of the second cluster. This indicates that airports with many runways have relatively low delay in the morning. This suggests that airports with more runways have relatively lower morning delay. A possible explanation for this result may be that airports with 1 or 2 runways have at most 1 runway available full time for departing aircraft. If a delay occurs, it will influence all departing flights. If more than three runways are present, and a problem occurs causing a delay, flights that are not related to this problem can divert to another runway.

B. Costs of Delay

Now that a broad picture is sketched regarding the performance of types of airports, the performance of individual airports and airlines can be discussed. Specifically, the extent of delay and its variation can serve as good performance indicator for airlines and airports. To get a better picture of the effect of delays, it was also of great importance to compute the financial cost of the delays.

Table 6 shows the five worst performing airlines, in terms of average delay over a day, and Table 7 shows the five best. The used nomenclature for the airlines can be found in Table 10 in Appendix A. It is noteworthy that the delay in Table 7 and Table 6 represents the delay per flight, and the cost represents the cost per passenger of that airline. It should be noted that for Table 7, HBJ is used for flights with an undefined airline.

Table 6. The five worst performing airlines ranked based on average delay per flight.

Rank	ICAO	Name	Delay [min]	σ_{delay} [min]	Cost [€]	σ_{cost} [€]
25	LLC	Small Planet Airlines	17.500	7.143	3.308	1.617
24	AUR	Aurigny Air Services	17.113	6.014	3.154	1.33
23	TCX	Thomas Cook Airlines	16.579	6.537	2.887	1.438
22	LOG	Loganair	16.186	7.082	2.764	1.256
21	TOM	Thomsonfly	16.084	5.817	2.851	1.155

Three of the five worst performing airlines (Small Planet Airlines, Thomas Cook Airlines, and Thomsonfly), are focused on leisure and holiday travel. These airlines aim to minimize their costs to provide cheap travel and holidays. Though it may then seem controversial that their cost is the highest, this could be because they buy cheaper time slots that are more frequently subjected to delays. Another possible explanation is that they have a lower priority for taking off when the airport is experiencing delays [16]. The other airlines, Aurigny Air Services and Loganair, are airlines operating in British airspace. As was stated before, this could also indicate a problem in British air traffic management. For the best performing airlines two are Swiss: Darwin Airline and Edelweiss Air. Daimler Chrysler Aviation is an airline company that operates smaller jets; the take-off requirements and boarding time for these types of aircraft is much lower than for larger jets, and results in a lower take-off delay. The final airline in the top 5 is Dniproavia, which operates mainly in Ukraine, East-European countries, and Austria [17]. The political tension between Russia, Europe and Ukraine may cause this airspace to be less crowded [18].

Table 7. The five best performing airlines ranked based on average delay per flight.

Rank	ICAO	Name	Delay [min]	σ_{delay} [min]	Cost [€]	σ_{cost} [€]
1	DCS	Daimler Chrysler Aviation	2.516	1.460	0.072	0.111
2	HBJ	Undefined	3.033	1.891	0.239	0.494
3	CSA	Czech Airlines	3.527	0.437	0.346	0.073
4	DWT	Darwin Airline	4.100	0.797	0.161	0.117
5	EDW	Edelweiss Air	4.264	1.821	0.219	0.143

The low standard deviation in addition to the low average delay per flight and average cost per passenger, are highly desirable properties for an airport. Not only is the delay cost typically low on average for airlines operating in these airports, but it is relatively stable in comparison to the deviation in delay times that can occur at other airports that were investigated.

With the aim of performing a similar analysis to that applied for airlines, the average delay per flight and average cost per passenger for each airport could be graphically represented as shown in Fig. 7. The vertical bars correspond to the standard deviation of the respective values, and the dashed lines indicate the averages of the entire data set for cost per passenger and delay per flight respectively.

Additionally, one can construct a similar graphical representation as shown in Fig. 7 for airlines. This figure can be found in Appendix A. The airlines are not individually labelled, but rather indicated with a

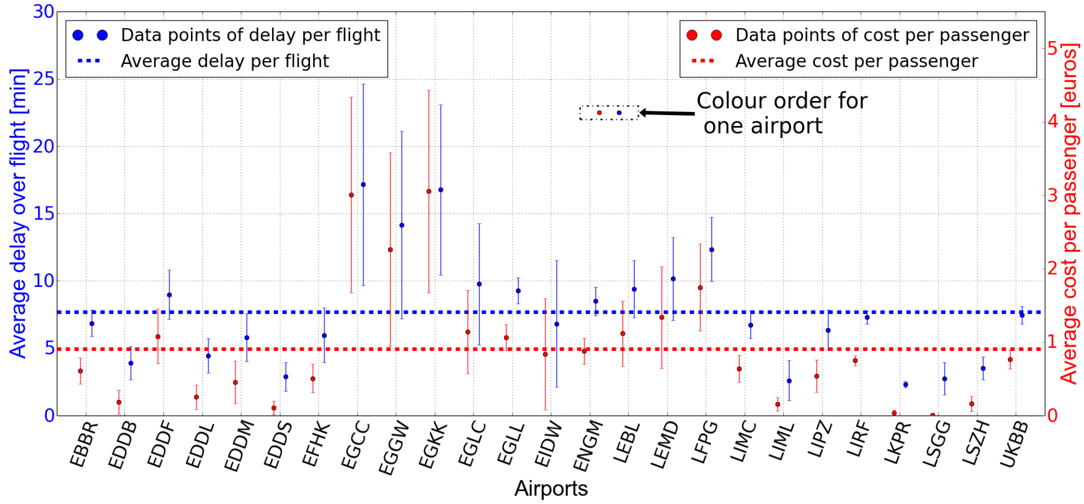

Figure 7. Graphical representation of mean departure delay and cost, and standard deviation for the airports in Table 1.

number; 1 to 103. Furthermore, the five worst performing airports (in terms of average delay per flight) could be tabulated as done in Table 9; the five best performing airports are tabulated in Table 8.

Table 8. The 5 worst performing airports ranked in terms of delay.

Rank	ICAO	Name	Delay [min]	σ_{delay} [min]	Cost [€]	σ_{cost} [€]
25	EGCC	Manchester	17.161	7.484	3.005	1.333
24	EGKK	London Gatwick	16.774	6.332	3.053	1.38
23	EGGW	London Lutton	14.148	4.505	2.261	1.32
22	LFPG	Paris-Charles de Gaulle	12.345	2.372	1.747	0.591
21	LEMD	Adolfo Suarez Madrid-Barajas	10.152	3.091	1.339	0.69

The significant deviation in the average delay per flight in the worst airports indicates that not only is the average delay high, but the value of this delay varies greatly from day to day; making this a variable difficult to predict. These airports are therefore highly unattractive for airlines in terms of the associated delay costs, taking into account that a few high-delayed flights are more expensive than many low-delayed flights (one 20 minute delay is worse than 20 one minute delays).

Furthermore, the 3 worst airports are located in the United Kingdom, of which 2 are located in London. This is analogous to the result found in Sec. IV-A.

Table 9. The 5 best performing airports ranked in terms of delay.

Rank	ICAO	Name	Delay [min]	σ_{delay} [min]	Cost [€]	σ_{cost} [€]
1	LKPR	Prague Vaclav Havel	2.326	0.242	0.036	0.039
2	LIML	Milan Linate	2.605	1.482	0.154	0.09
3	LSGG	Geneva	2.736	1.208	0.005	0.005
4	EDDS	Stuttgart	2.892	1.059	0.105	0.092
5	LSZH	Zurich	3.528	0.856	0.16	0.102

The five best performing airports have a low mean delay; in addition to having a small standard deviation. In fact, it can be concluded with 68% certainty that the average delay per flight at these airports will be less than five minutes. The Vaclav Havel Airport Prague is the airport with the lowest delay and also has a very small standard deviation. The average delay on this airport is 2.326 minutes which makes it the best performing airport from the remaining data set in terms of delay. Some flights are still delayed more than five minutes at LKPR, resulting in a finite cost. At Geneva Airport (LSGG), almost no aircraft were delayed more than five minutes, resulting in an almost zero cost (since delays less than 5 minutes are associated with no cost as explained previously). However, the average delay and standard deviation for LSGG were still higher than for LKPR.

C. Relation between Morning Delay and Delay Costs

Now that the performance of airports regarding morning delay and their costs have been discussed, it can be evaluated whether there is a relationship.

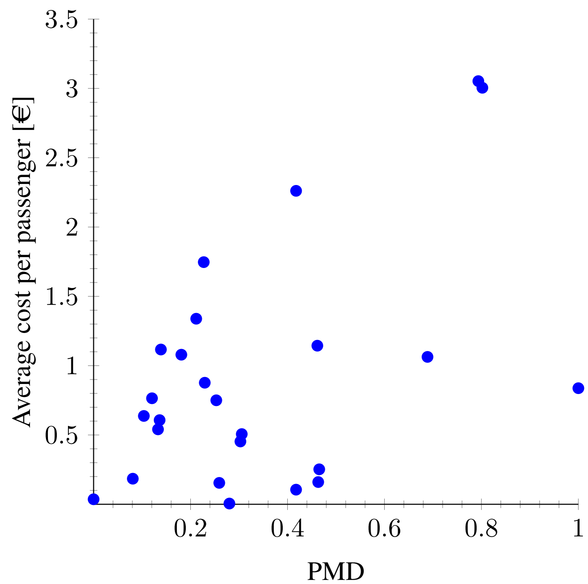

Figure 8. Airports grouped on size of airport.

In Fig. 8, the average cost per passenger over the entire day is plotted against the performance parameter PMD explained in Sec. III-A. From inspection, it is apparent that the worse the performance in mornings (high average delay during the morning and low ΔT), indicated by a higher PMD, lead to higher average costs per passenger throughout the entire day. This is confirmed by the correlation coefficient, which equals $\rho = 0.493$; this correlation is statistically significant, $p < 0.05$ [19].

This is not a trivial result: apparently, airports that suffer badly from delays in the morning are more likely to suffer from high costs associated with delays, meaning these morning delays show propagating behaviour throughout the day. It should be noted, however, that there is a general lack of airports with a relatively high PMD. This means that the possible presence of outliers can greatly affect the correlation coefficient. Consider London Gatwick and London Lutton airport, which both have extraordinary high costs per passenger; discarding these data points results in a correlation coefficient of $\rho = 0.141$; this correlation is statistically insignificant, $p > 0.05$ [19]. As the sample size is too small to reliably determine whether these airports are indeed outliers or not, further research should aim to obtain a larger amount of data to determine whether these are outliers or not.

V. Conclusion and Recommendations

As could be seen from Table 4 and Table 5, the best and worst airports in terms of the PMD are Prague Vaclav Havel and Dublin, respectively. It becomes apparent from Table 5 that four of the five worst airports are located on the British isles, which could be due to an overly crowded airspace, particularly above the city of London. It was shown that there is no statistically significant relationship between airport size, and severity and frequency of morning delay. However, airports with three or more runways were shown to have a tendency to have lower and less severe morning delay. This is likely due to the fact that having more runways allows flights to be diverted should problems occur.

From the performance and cost analysis of both airports and airlines, it was evident that a very specific set of airports and airlines have significantly more frequent delay occurrences and high associated costs than others. The most notable examples being Manchester and London Gatwick airport, and Aurigny Air Services and Small Planet Airlines. Not only this, but these airports and airlines typically displayed a much higher variation in both delay and associated hard costs. This means that these airlines perform in a highly undesirable manner, as flights from this airline not only experience high magnitudes of delays in general, but also show a large variation in these delays and costs. Likewise, airports with these characteristics are highly undesirable for airlines both due to the high average delay and cost for flights departing from the airport and the large degree of variation (unpredictability) in both of these aspects.

Finally, in Sec IV-C it was found that there is a statistically significant correlation of 0.493 between the PMD and the average cost per passenger on an airport. This implies that airports with a high frequency of severe delays in the morning tend to have higher costs associated with delay over the entire day. This can be explained by the fact that delay propagates over the entire day once severe delays have occurred. The relationship between PMD and the average cost per passenger on an airport should be put into perspective however. The value of the correlation coefficient is not entirely reliable as the value is greatly dependant on the presence of outliers. However, outliers can not be identified as the sample size of twenty-five airports is too small. For further research, it is preferred to use a more advanced data set. Due to the quality of the data provided by Eurocontol, it was required to filter 76.53% of all flights. As a result, flight delay patterns in Europe could not be analysed fully and sample sizes were too small to establish reliable links. Additionally, more information regarding the flights would be useful. Important points missing in the data set were arrival delay and tail number. This information could be used to track aircraft during the day, allowing to investigate how delays propagate over Europe. Furthermore, research into soft costs would be interesting, as it would allow for an even better representation of the costs caused by flight delays in Europe.

VI. Acknowledgments

We would like to thank dr. ir. Bruno F. Lopes dos Santos for his help and guidance during the research, as well as the Delft University of Technology for giving us this unique research opportunity. Second, we thank Drs. Angeniet Kam for the useful tips and corrections on the writing style of this paper. Lastly, we would like to thank Eurocontrol for the valuable aircraft data they sent us for analyzing.

References

[1] IATA, *IATA Forecasts Passenger Demand to Double Over 20 Years*, URL http://www.iata.org/pressroom/pr/Pages/2016-10-18-02.aspx, [Accessed June 9 2017].

[2] European commission of mobility and transport, *Airport capacity and quality*, URL https://ec.europa.eu/transport/modes/air/airports/airport_capacity_and_quality_en, [Accessed June 9 2017].

[3] Eurocontrol, *Delays to air transport in Europe - Annual 2013*, URL https://www.eurocontrol.int/publications/coda-digest-annual-2013, [Accessed May 18 2017].

[4] Institut du Transport Aerien, *Cost of air transport delay in Europe*, 2000.

[5] Jetzki, M., *The propagation of air transport delays in Europe*, Master's thesis, RWTH Aachen University, December 2009.

[6] Fleurquin, P., Ramasco, J. J., and Eguiluz, V. M., *Systemic delay propagation in the US airport network*, arXiv preprint arXiv:1301.1136, 2013.

[7] Cook, A., Tanner, G., and Lawes, A., *The Hidden Cost of Airline Unpunctuality*, Journal of Transport Economics and Policy, May 2012, vol. 46, pp. 157 – 173.

[8] Cook, A., Tanner, G., Jovanovi, R., et al., *The cost of delay to air transport in Europe Quantification and management*, Air Transport Research Society (ATRS) World Conference, June 2009, vol. 13.

[9] Eurocontrol, *Central Office for Delay Analysis*, URL https://www.eurocontrol.int/articles/central-office-delay-analysis-coda, [Accessed May 12 2017].

[10] Eurocontrol, *Our Member States, Partner Countries & Organisations*, URL https://www.eurocontrol.int/about/member-states, [Accessed June 13 2017].

[11] Federal Aviation Administration, *Airport Categories*, URL https://www.faa.gov/airports/planning_capacity/passenger_allcargo_stats/categories/, [Accessed June 8 2017].

[12] Maimon, O., and Rokach, L., *Data Mining and Knowledge Discovery Handbook*, 2 ed., Springer, 2010, pp. 332-334.

[13] Field, A., *Discovering Statistics Using SPSS*, third ed., Sage, 2009.

[14] van der Duim, T., van Elsloo, S., Haegens, T., et al., *Literature Review on Aircraft Delay in Europe*, March 2017, pp. 1 – 6.

[15] Eurocontrol, *Standard Inputs for Eurocontrol Cost-Benefit Analyses*, URL https://www.eurocontrol.int/sites/default/files/publication/files/standard-input-for-eurocontrol-cost-benefit-analyses-2015.pdf, [Accessed May 12 2017].

[16] Eurocontrol, *Planning for Delay: inuence of ight scheduling on airline punctuality*, URL https://www.eurocontrol.int/download/publication/node-field_download-4686-0, [Accessed June 13 2017].

[17] Dniproavia, *Dniproavia*, URL http://dniproavia.com/en/, [Accessed June 10 2017].

[18] Harper, K., *Ukraine bans Russian airlines, strengthening tensions with Moscow*, The Washington Post, July 2015, URL https://www.washingtonpost.com/world/europe/ukraine-bans-russian-airlines-in-new-twist-to-showdow-with-moscow/2015/09/25/3f0fd3cb-1519-4038-bcb0-5adc5ec662a7_story.html?utm_term=.3bcbc2742d68, [Accessed June 10 2017].

[19] Weathington, B. L., Cunningham, C. J. L., and Pittenger, D. J., *Understanding Businnes Research*, chap. Appendix B: Statistical Tables, John Wiley & Sons, Inc., 2012, doi: 10.1002/9781118342978.app2.

Appendix A

This appendix contains plots showing the average delay per flight and the average cost per passenger (Fig. 9 and 10) of all airlines considered in the data set. A legend of all airlines considered is shown in Table 10.

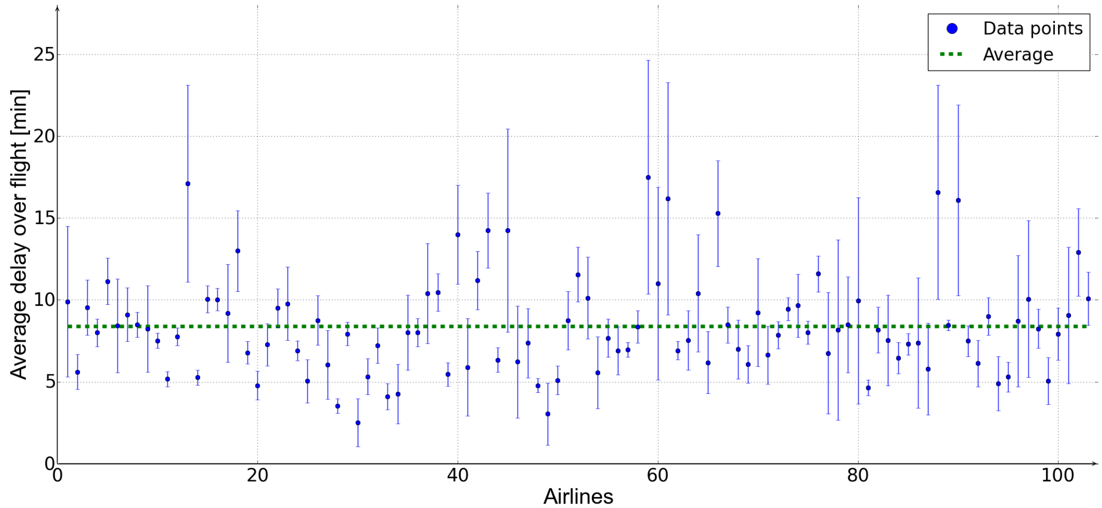

Figure 9. Average delay per passenger and standard deviation for the airlines in Table 10.

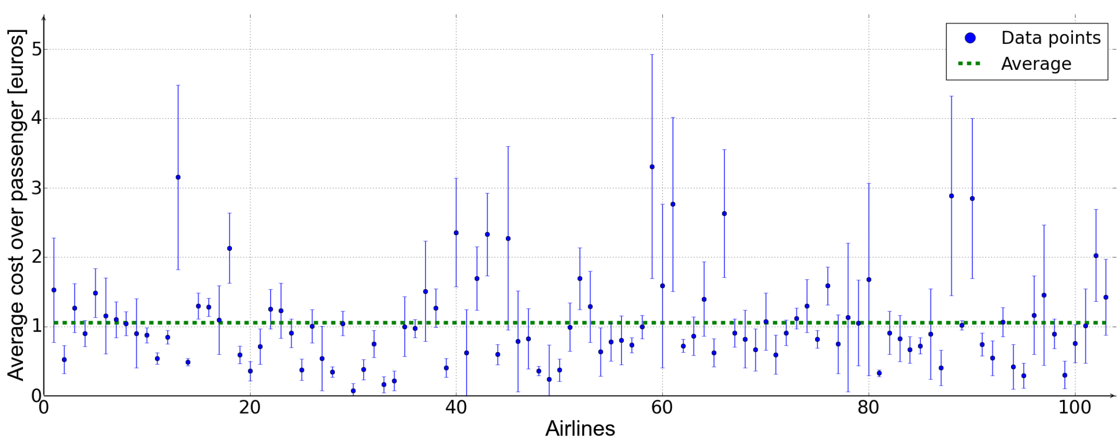

Figure 10. Average cost per passenger and standard deviation for the airlines in Table 10.

Table 10. Legend of all airlines used in this paper.

No.	ICAO	Airline	No.	ICAO	Airline
1	ABR	ASL Airlines Ireland	53	IBS	Iberia Express
2	ADR	Adria Airways	54	ISS	Meridiana
3	AEA	Air Europa	55	JAF	Jetairfly
4	AEE	Aegean Airlines	56	KKK	Atlasglobal
5	AFR	Air France	57	KLM	KLM Royal Dutch Airlines
6	AHO	Air Hamburg	58	LGL	Luxair
7	AHY	Azerbaijan Airlines	59	LLC	Small Planet Airlines
8	AMC	Air Malta	60	LNX	London Executive Aviation
9	ANE	Air Nostrum	61	LOG	Loganair
10	ASL	Air Serbia	62	LOT	LOT Polish Airlines
11	AUA	Austrian Airlines	63	LZB	Bulgaria Air
12	AUI	Ukraine Intl. Airlines	64	MGX	Montenegro Airlines
13	AUR	Aurigny Air Services	65	MLD	Air Moldova
14	AZA	Alitalia	66	MON	Monarch Airlines
15	BAW	British Airways	67	NAX	Norwegian Air Shuttle
16	BCS	European Air Transport	68	NJE	NetJets Europe
17	BCY	Cityjet	69	NLY	Niki
18	BEE	Flybe	70	NOS	Neos
19	BEL	Brussels Airlines	71	OHY	Onur Airlines
20	BER	Air Berlin	72	PGT	Pegasus Airlines
21	BMR	BMI Regional	73	ROT	Romanian Air Transport
22	BMS	Blue Air	74	RYR	Ryanair
23	BPA	Blue Panorama Airlines	75	SAS	Scandinavian Airlines
24	BTI	Air Baltic	76	SHT	Sepehran Airlines
25	CAI	Corendon Airlines	77	SPR	Provincial Airlines
26	CFE	Cityflyer Express	78	STK	Stobart Air
27	CFG	Condor Flugdienst	79	SUS	Sun Air of Scandinavia
28	CSA	Czech Airlines	80	SWN	West Air Sweden
29	CTN	Croatia Airlines	81	SWR	Swiss Intl. Air Lines
30	DCS	Daimler Chrysler Aviation	82	SWT	Sunworld Airlines
31	DLA	Air Dolomiti	83	SXD	SunExpress Deutschland
32	DLH	Lufthansa	84	SXS	SunExpress
33	DWT	Darwin Airline	85	TAP	TAP Portugal
34	EDW	Edelweiss Air	86	TAY	TNT Airlways
35	EIN	Air Lingus	87	TCW	Thomas Cook Airlines Belgium
36	ELY	El Al	88	TCX	Thomas Cook Airlines
37	ENT	Enter Air	89	THY	Turkish Airlines
38	ETH	Ethiopian Airlines	90	TOM	Thomsonfly
39	EWG	Eurowings	91	TRA	Transavia
40	EXS	Jet2	92	TUI	TUI Fly
41	EZS	EasyJet Switzerland	93	TVF	Transavia France
42	EZY	EasyJet	94	TVS	Travel Service
43	FDX	FedEx Express	95	UDN	Dniproavia
44	FIN	Finnair	96	UPS	United Parcel Service
45	FPO	Europe Airpost	97	VJT	Vistajet
46	GAC	GLobeAir	98	VLG	Vueling Airlines
47	GMI	Germania	99	VOE	Volotea
48	GWI	Germanwings	100	WIF	Wideroe
49	HBJ	Undefined	101	WRF	Wright Air Service
50	HOP	HOP!	102	WZZ	Wizz Air
51	IBE	Iberia	103	ZZZ	Undefined
52	IBK	Norwegian Air Intl.			

View publication stats