Statistical processing of arrivals and departures on Košice Airport in 2007

Štefan Berežný

This article describes the processing of records of flights from Košice airport and subsequent acquisition and processing of information contained in these data. The results of processed real data show the delays of flights in 2007.

1 INTRODUCTION

Statistical processing and evaluation of data, which are collected during creating records on individual flights to Košice airport, are important source of information regarding the operations and their parameters. Partial processing of these data has been shown in work [1]. Data were provided for the processing from Kosice airport Deputy Executive Director for the operation and safety at the airport Košice Mrs. Ing. M. Horváthová. Names of airline in this article are not real. (We use other names/markers, which are described later.) The substitution of airline names is the same as in paper [2] and also in article, which describes the processing of similar data of 2006 year. This article refills the missing data between 2006 and 2008 year. The completeness of these data is necessary for further analysis, which we will come up in the conclusion of this article.

For the processing and evaluation of the data we used software Excel from Microsoft Office software package, the Matlab 7 and QOctave.

The entire procedure can be summarized in three phases. The first stage includes the preparing the data to further processed using statistical and mathematical functions, which provides the described software. In the second phase of the process we compile the adjusted data. In the third stage, the interpretation of the data was obtained in the way of the assumptions that were established in the first and second phase.

This article is also built on the article [2] in which the data of 2008 are processed and article, which describes the data processing from Košice International Airport in 2006. By this we create a time-series data which can be analyzed in time later or comparing the above parameters in individual years or months. In addition to the processed data the analysis of year 2007 are available. So we gain the database with a complete interval from 01. 01. 2006 to 31. 12. 2008. Such a database provides enough information for time analysis and comparison of the relevant time periods in different seasons. These analyses are in further examination of the abovementioned processing and evaluation of the results obtained from the data provided.

2 DATA PROCESSING AND PRESENTATION

Provided data include also the following characteristics:

DATE - the date, when record has been made,
CARIERE - the name of the airline, which took the listed flight
FLYGHT NUMBER - flight number,
ARRIVAL - Boolean value that determines if this flight was the arrival flight,
DEPARTURES - Boolean value that specifies if this flight was the departure flight,
STM - time of arrival, respectively departure, which is given in hours and minutes,
ATM - the actual time of arrival, respectively departure, which is also given in hours and minutes,

DL - delay of the aircraft (flight) in minutes.

In database are 9 500 records were checked from the time interval from 01. 01. 2007 to 31 12. 2007.

Due to the nature of the data it was necessary to adapt these data for further processing and evaluation. Editing of input data was necessary in view of the used functions in excel and MATLAB.

The first step which we did, it was the removing ambiguities in the identification of all the data of flights. It was necessary to carry out an inspection of all records provided by the missing data, or adjust existing data to the information was not lost and that it was possible to identify all the necessary data to the flight, which we need for further calculations and processing.

The second step includes a visual adjustment of the data in order to move easily in such data. After such adjustments we could create formulas easier which are necessary for carrying out the analysis and thus also eliminates the risk of errors in creating patterns and functional dependencies.

3 DATA ANALYSIS

During analyzing the data, we focused mainly on analyzing the temporal time deviation that arose in the execution of each recorded flight to delay the planned range and arrivals at the airport Košice.

These deviations for each flight, we had to calculate firstly and enter the additional data to existing dataset for all records in that period. If it was a delay (i.e., the aircraft landed respectively took off later, as the flight was scheduled), we entered the time delay with a positive sign. If flight took off or landed earlier before its scheduled arrival (departure) time, we entered that time difference for all records with a negative sign. After the creation of additional records we had to identify the flights, which have not been realized. These records had to be omitted and do not use them for further analysis.

Due to the limited space in this article we show only selected statistics, which provide a relevant image of the monitored parameters and features of flights on Kosice airport.

Here are the following results:

- The expected value of time deviation from the scheduled time of arrival respectively departure in minutes.
- The expected value of time deviation from the planned time of arrival.
- The expected value of time deviation from the planned time of departure.
- The expected value of time deviation from the scheduled time of arrival, respectively departure, with respect to flights made by individual airlines.
- The expected value of time deviation from the scheduled time of arrival for flights made by the individual airlines.
- The expected value of time deviation from the planned departure time for flights made by individual airlines.
- The expected value of time deviation from the scheduled time of arrival respectively departure in view of the day of the week.
- The expected value of time deviation from the scheduled time of arrival respectively departure in respect of the current month.
- The expected value of time deviation from the scheduled time of arrival respectively departure with respect to flights divided on scheduled flights and charter flights.
- Variances and standard deviations for the samples.

In next part we marked the airlines by symbols AIR-01 - AIR-16, when first 14 marks mean the airlines with biggest market share on Košice Airport and AIR-16 means the all other airlines with smaller market share. These airlines usually do not fly to Košice periodically. The special group of flights was marked as AIR-15.

In table (Tab. 1) is clearly shown the usage of airline marks in Košice International Airport for this article, in article [2] and in article, in which are processed data of 2006 year. The marks in the same rows mean the same airlines.

In table (Tab. 2) and graph (Graph 2) we show the number and share of each airlines on realized arrivals and departures on Košice airport.

In this table 171 records are not included, which means do not realized flights or it means flights, for which was not checked time of arrival

or departure and also planned arrival and departure time. All these records were being excluded from next analysis and published values are adjusted from these records.

We had to refill other 7 auxiliary characteristics to existing 40 characteristics in every record (row), which helps us for next analysis and data processing. By these changes we got base set, which contains 9089 rows and 47 columns with airlines marked as AIR-01 to AIR-15. Other airlines included to mark AIR-16 have been checked in 231 rows of records.

Airlines in 2008	Airlines in 2007	Airlines in 2006
AIR-01	AIR-01	AIR-01
AIR-02	AIR-02	AIR-02
AIR-03	-	-
AIR-04	AIR-03	AIR-03
AIR-05	AIR-04	AIR-04
AIR-06	AIR-05	AIR-05
AIR-07	-	-
AIR-08	AIR-06	AIR-06
AIR-09	-	-
-	AIR-07	-
AIR-10	AIR-08	AIR-07
AIR-11	-	-
AIR-12	-	-
-	-	AIR-08
AIR-13	AIR-09	AIR-09
AIR-14	AIR-10	AIR-10
-	-	AIR-11
AIR-15	AIR-11	AIR-12
-	AIR-12	AIR-13
AIR-16	AIR-13	AIR-14
AIR-17	-	-
AIR-18	AIR-14	AIR-15
AIR-19	AIR-15	AIR-16
AIR-20	AIR-16	AIR-17

Tab. 1: An overview of the corresponding names of the airlines in this article and in article [2] *and in article, in which are processed the data of 2006*

The reader easily notices that the greatest share on Košice Airport had the airline AIR-05 up to 36 %. After it there are the AIR-06 with share of 21 % and AIR-02 with share of 12 %. The airline AIR-11 had the share of 5 %.

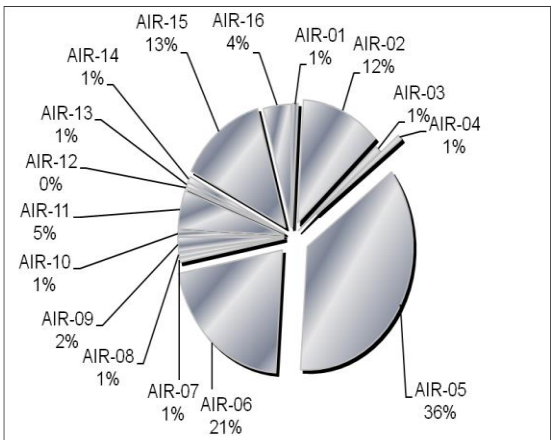
Airlines	Nº of Arr	Nº of Dep	SUM
AIR-01	40	40	80
AIR-02	572	569	1141
AIR-03	42	42	84
AIR-04	33	32	65
AIR-05	1703	1700	3403
AIR-06	978	980	1958
AIR-07	28	28	56
AIR-08	30	30	60
AIR-09	83	83	166
AIR-10	34	37	71
AIR-11	235	235	470
AIR-12	15	15	30
AIR-13	26	26	52
AIR-14	38	38	76
AIR-15	611	609	1220
AIR-16	200	197	397
SUM	4668	4661	9329

Tab. 2: Number of realized arrivals and departures by airlines on Košice airport

In table (Tab. 3) and graph (Graph 3) we show the expected value of time deviation from the planned time of arrival, time of departure and time of arrival respectively departure in minutes.

Airlines	Arrival	Departure	Arr + Dep
AIR-01	25,83333	26,79167	26,3125
AIR-02	0,493852	3,807377	2,150615
AIR-03	24,0625	18,3125	21,1875
AIR-04	-2,22917	1,145833	-0,54167
AIR-05	1,4	19	10,2
AIR-06	5,067882	-1,21195	1,928953
AIR-07	18,34524	22,52381	20,43452
AIR-08	-1,96471	9,377918	3,703026
AIR-09	14,79747	14,53165	14,66456
AIR-10	40,42553	26,04255	33,23404
AIR-11	21,4375	-11	6,3
AIR-12	12,82432	18,73333	15,79866
AIR-13	3,818182	13,45455	8,636364
AIR-14	9,368421	27,55556	18,21622
AIR-15	27,06436	38,09453	32,56576
AIR-16	9,68	10,32	10
TOTAL	6,066255	8,395277	7,234071

Tab. 3: Number of realized arrivals and departures by each airlines on Košice Airport



Graph 3: Time deviation

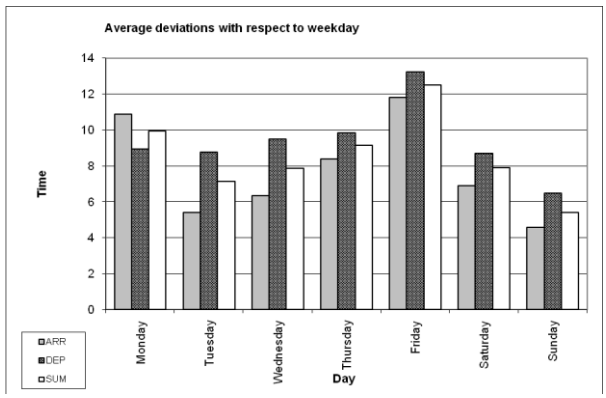
The average delay of flight on Košice airport is over 7 minutes, but arrivals delay about 6 minutes and departures delay over 8 minutes. In table (Tab. 3) is shown also the 5 negative values, which mean the sooner arrival or departure in relation to schedule time. Ideal values are the differences nearby zero. We checked that only AIR-04 has the negative total value of delay, so mostly only few flights are delay. It is ideal for the passengers because they know that they do not miss the flight and to finish airport they come earlier than it is planned. Bits of unusual are the values of AIR-11, but without deeper analysis we cannot identify their causes.

Next 2 tables and graphs show the time deviation of arrival and departure time by the day of week and month in year.

In table (Tab. 4) and subsistent graph (Graph 4) there is survey of flight delay in respect to day in week.

Day	Arrival	Departure	Arr + Dep
Monday	10,89	8,93	9,94
Tuesday	5,40	8,77	7,13
Wednesday	6,36	9,50	7,89
Thursday	8,40	9,85	9,14
Friday	11,80	13,25	12,53
Saturday	6,90	8,69	7,92
Sunday	4,58	6,49	5,42
TOTAL	8,10	9,69	8,89

Tab. 4: Average deviations regarding weekday



Graph 4: Average deviations with respect to weekday

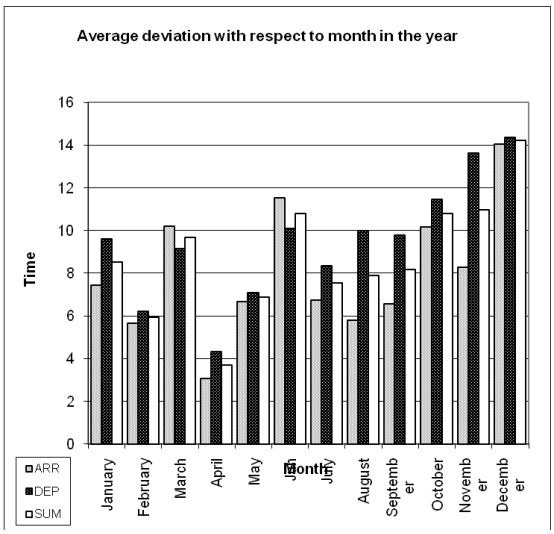
In table (Tab. 5) and graph (Graph 5) we show the survey of flight delay in respect to month of year.

Months	Arrival	Departure	Arr + Dep
January	7,44	9,61	8,52
February	5,64	6,22	5,93
March	10,21	9,15	9,68
April	3,05	4,31	3,68
May	6,68	7,101	6,89
Jun	11,53	10,08	10,79
July	6,75	8,36	7,55
August	5,80	9,98	7,89
September	6,56	9,78	8,18
October	10,15	11,44	10,79
November	8,29	13,64	10,97
December	14,05	14,35	14,20
TOTAL	8,10	9,69	8,89

Tab. 5: Average deviation regarding month in the year

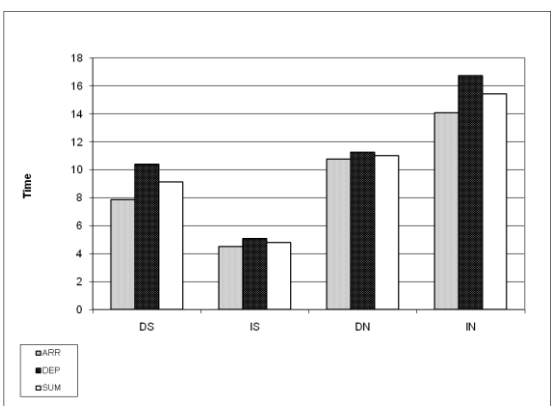
Type of flight	Arrival	Departure	Arr + Dep
Domestic Schedule (DS)	7,85	10,38	9,12
International Schedule (IS)	4,52	5,09	4,81
Domestic nonschedule (DN)	10,75	11,25	10,99
International nonschedule (IN)	14,08	16,72	15,41
TOTAL	8,10	9,69	8,89

Tab. 6: The time deviation in respect to Schedule and no schedule flights and international and domestic flights



Graph 5: Average deviation with respect to month in the year

In table (Tab. 6) and graph (Graph 6) we show the time deviation in respect to schedule and no schedule flights and international and domestic flights.



Graph 6: The time deviation in respect to Schedule and nonschedule flights and international and domestic flights

4 ANALYSIS OF RESULTS

The results in the third chapter give us an initial image how the time deviations are depended on the different measured characteristics. Based on the processed data and graph we can choose the next steps of these data analysis. Some applications of results are shown in [6].

Similar analysis of data of flights we can find also for airports Chicago O Hara a Schiphol [4] [5]. These publications include similar approach for data analysis, but the results are adequate to airport size, which they describe.

In additional analysis of these data we can use the Six Sigma method and TQM, which it should be done by enclosed source [7] [8] [9]. Economic impact analysis of flights is very important in additional analysis. [10].

Similar problems of data processing as in our case had been addressed in the analysis of fractures of conveyor belts, where the issue of data processing is described in articles [11] [12].

From processed data we can see that positive deviation preponderates before negative deviation, which mean to delay and earlier arrivals and departures. The strong difference we can see in case of analysis data in respect of type of flights. The Schedule flight indicates more the negative deviations before positive deviations. This parameter could be analyzed in detailly. In this paper we do not have enough places for this addition analysis.

Thus, the processed data are the basis for further analysis. One of other possibilities is the analysis of time series that was generated during data processing. A suitable method of processing this type data is described in the work [13] [14].

The methods of other analysis of obtained data are described in work [15] to gain the other characteristics for evaluation of time data.

REFERENCES

[1] LACZKÓ Radoslav: Štatistické spracovanie meškaní letov na Košickom letisku. Bakalárska práca. Košice: Technická univerzita v Košiciach, Letecká fakulta, 2009. 48 s.

[2] BEREŽNÝ Štefan, ANDREJKOVIČ Marek: Statistical processing of flights on Košice airport, Acta Avionica. - ISSN 1335-9479. - Roč. 11, č. 17 (2009), s. 9-14.

[3] BEREŽNÝ Štefan - ANDREJKOVIČ Marek The statistical outcomes in the airport as a tool of decision-making and management process, Liberec Economic Forum 2009 : proceedings of the 9th international conference : Liberec, 15th - 16th September

2009. - Liberec: Technical University, 2009. - ISBN 978-80-7372-523-5. - P. 52-58.
[4] MC MILLEN, D.: Airport expansions and property values: The case of Chicago O Hare airport, Journal of Urban Economics 55, 627-640, 2004
[5] MORRELL, P., LU C.: Aircraft noise social cost and charge mechanisms – A case study of Amsterdam airport Schiphol, Transportation Research Part D 5, 305-320, 2000.
[6] TKÁČ, M. – ANDREJKOVIČ, M. – HAJDUOVÁ, Z.: Analysis of Sky Europe airlines flights on Košice international airport. In: *Acta avionica*. - Košice : Technical University Kosice, Faculty of aeronautics, 2009. - ISSN 1335-9479. - Vol. XI, č. 17 (2009) s. 174-179
[7] PIĽA, J. – HAJDUOVÁ, Z. – ANDREJKOVIČ, M.: Aplikácia metódy ANOVA v strojárskom priemysle. In: *Bezpečnosť - kvalita - spoľahlivosť* : 4. medzinárodná vedecká konferencia, Košice 2009 : zborník z konferencie / Recenzovali: Aurel Sloboda a kol. - Košice : Technická univerzita, Strojnícka fakulta, Katedra bezpečnosti a kvality produkcie, 2009. - ISBN 978-80-8073-828-0. - s. 178-183
[8] ANDREJKOVIČ, M. – HAJDUOVÁ, Z.: Analýza reklamácií v procesoch opráv = Analysis of claims in the processes of repair. In: Forum statisticum Slovacum : vedecký časopis Slovenskej štatistickej a demografickej spoločnosti. - Bratislava : Slovenská štatistická a demografická spoločnosť, 2009. - ISSN 1336-7420. - Roč. 5, č. 2 (2009), s. 159-163
[9] TURISOVÁ, Renáta - HAJDUOVÁ, Zuzana: Nástroje Six Sigma pri riadení leteckej prevádzky. In: Nové trendy rozvoja letectva : 7. medzinárodná vedecká konferencia, Košice, September 6 - 8, 2006 : Zborník príspevkov. Košice : TU, 2006. 7 s. ISBN 80-8073-520-4.
[10] ANDREJKOVIČ Marek: Odhady ekonomických parametrov pri hodnotení návratnosti projektov využívania obnoviteľných zdrojov energie. In *Acta Oeconomica Cassoviensia*, roč. I, 2008, č. 2, s. 28 – 35, ISSN 1336-6020
[11] BINDZÁR Peter - GRINČOVÁ Anna - RISTOVIC Ivica: 3D mathematical model of conveyor belt subjected to a stress loading, in: podzemni radovi. - ISSN 0354-2904. - vol. 13, no. 15 (2006), p. 179-188.
[12] GRINČOVÁ Anna - hlúbiková Adria - KREŠÁK Jozef: Metodika skúšania dopravných pásov pri prieraze, In: Doprava a logistika = = Transport & Logistics. - ISSN 1451-107X. - Mimoriadne č. 5 (2008), s. 209-213.
[13] KIMÁKOVÁ Zuzana - ANDREJIOVÁ Miriam: analýza časového radu emisií sox pomocou programu r. In: Forum statisticum slovacum. Nitra, 2009. 6/2009. Ročník v. Str. 63-68. ISSN 1336-7420.
[14] ANDREJIOVÁ Miriam - KIMÁKOVÁ Zuzana: Prognózovanie predpovedí v časových radoch pomocou exponenciálneho vyrovnávania a harmonickej analýzy. In: FORUM STATISTICUM SLOVACUM. Nitra, 2008. 4/2008. ročník IV, str. 8 -13. ISSN 1336-7420.
[15] SOTÁK Miloš: Estimation of stochastic coefficients of inertial sensors. In: Science & Military. - ISSN 1336-8885. - Roč. 3, č. 2 (2008), s. 13-16.

AUTHOR

RNDr. Štefan BEREŽNÝ, PhD.
Technical University in Košice
Fakulty of Electrical Engineering and Informatics

Department of Mathematics
B. Němcovej 32
042 00 Košice
Slovak republic
e-mail: Stefan.Berezny@tuke.sk
telefón: 055 602 2447
mobil: 0918 400 579