Journal of Air Transport Management 57 (2016) 1–6

Contents lists available at ScienceDirect

Journal of Air Transport Management

journal homepage: www.elsevier.com/locate/jairtraman

Note

Assessing performance variability of ground handlers to comply with airport quality standards

Riccardo Patriarca, Giulio Di Gravio, Francesco Costantino*

Department of Mechanical and Aerospace Engineering, Sapienza University of Rome, Via Eudossiana, 18, 00184, Rome, Italy

ARTICLE INFO

Article history:
Received 15 July 2015
Received in revised form
6 June 2016
Accepted 25 June 2016

Keywords:
Ground handling
Airport
Quality standard
Baggage delivery

ABSTRACT

The increasing competition in the air transport industry continuously pushes ground handlers to improve their performance. Recent European regulations prescribe that the managing body of large European airports (more than 5,000,000 passengers/year or 100,000 tonnes/year of freight) have to define quality standards of service level for ground handler operations. In case a handler fails to meet these minimum requirements, the airport has to report it these irregularities, potentially imposing a fine or even suspending, partly or fully, the handler's services. As always recognized by the passengers, one of the most critical quality standards is the baggage handling waiting time. This paper defines a methodology to assess whether the handler fits the requirements of first baggage delivery for both the overall process and the specific segments (flight, conveyor belt, departing airport, day of the week, etc.). The methodology, analysing the performance by a two-dimensional perspective (structure and frequency), will help the decision makers in defining mitigating actions to take the performance under statistical control, respecting the requirements and preventing any drift.

1. Introduction

Ground handling covers a wide range of services that support airline air operations. It may include both technical services, such as maintenance, fuel & oil services, freight handling, and passengers' essential services, i.e. check-in, disembarkation, surface transport, baggage handling, comfort.

This industry experienced a very dynamic development and growth, since the introduction of Council Directive 96/67/EC (EU, 1996). In accordance with the liberalisation of the air transport, the Directive aims at developing the competition among players to increase the efficiency of the ground handling activities, to decrease the average costs, increase the quality levels of services and enhance the choice for airlines (Soames, 1997). The definition of quality, which has to include safety matters, baggage delays, environmental matters, etc., evolves in a two-party service level agreement between airlines and ground handlers.

After the Directive's implementation, Airport Research Center (2009) showed a decreasing trend on prices and an increase of competition, given the large number of third party handling companies entering the market. Anyway, statistics on quality level were unclear. While many airport managing bodies and airlines are generally satisfied, others observe that quality levels suffered from this competition on costs (Costantino et al., 2013).

Hence, in 2013, the European Parliament (EU, 2013) asked for a development of ground handling quality levels to safeguard airports and airlines operations (Burghouwt et al., 2014). The regulation requires that large European airports with more than 5,000,000 passengers/year or 100,000 tonnes/year of freight for at least the previous three years set minimum quality standards to respect. These standards are (e.g.) maximum boarding and disembarkation time, maximum time for delivery of first and last baggage, minimum training level, maximum time for de-icing an aircraft, etc. Airport managing body has the right to impose a fine, restrict or prohibit a ground handler from supplying services if that supplier fails to comply with the requirements. Monitoring service performance acquires a fundamental role in airport management (Bezerra and Gomes, 2015; Chen and Chang, 2005).

This paper defines a methodology to assess the handler quality level in accordance with the minimum standards set for the operations, e.g. the maximum waiting time for delivery of first baggage. It allows verifying if the handler fits the standards and highlighting potential critical segments that, even if in compliance with the standards, may require early mitigating actions. The methodology

* Corresponding author.
E-mail address: francesco.costantino@uniroma1.it (F. Costantino).

http://dx.doi.org/10.1016/j.jairtraman.2016.06.009
0969-6997/

is easily applicable to other kind of performance as for last baggage delivery time and several other processes as well (waiting time for baggage check-in, time for passenger boarding/disembarkation, time for transferring passengers between connecting flights, etc.).

2. Method

As stated in Annex 1a, the EU Directive prescribes that the airport managing body shall set a maximum time for delivery of first and last item of baggage, for the airport as a whole or for an individual terminal (EU, 2013). The airport information system collects data about the block-on time of each landing aircraft and the delivery of first and last baggage. These data are made available to the handler's decision maker, which can compare them with the one reported by its operators. The following 4-step strategy (segmentation, structure analysis, frequency analysis, criticality analysis) can be applied to any of the two databases to assess the first baggage delivery time, with respect to the standards.

2.1. Segmentation

This stage consists of classifying the data of the information system (block-on time and first baggage delivery time). Although the relevance of segmentation can differ from case to case, according to the characteristics of the specific handler and airport, main variables to consider are, for example, airline, flight number, departure airport, aircraft and baggage conveyor belt. The delivery time in minutes is then the difference between these two terms.

2.2. Structure analysis

This stage consists of 3 sub steps:

- *Target Evaluation.* The first baggage delivery time (T_1), is a continuous random variable to represent with a probability density function (PDF). In detail, rather than T_1, it is possible to evaluate the PDF that computes the gap between the first baggage delivery time and the relative target time T_1^*, i.e. $\Delta T_1 = T_1 - T_1^*$. Naming L_1^* the number of first baggage to process within a target waiting time (on 100 occasions), the L_1^{*th} percentile of ΔT_1, i.e. $L_1^{*th}perc(\Delta T_1)$, describes the performance. $L_1^{*th}perc(\Delta T_1) \leq 0$ represents a segment which complies with the requirements, while $L_1^{*th}perc(\Delta T_1) > 0$ a process out of specification.
- *PDF Shape.* The normal distribution represents the performance variability under statistical control (no systematic cause acting). The gap between target delivery time and real delivery time for an ideal process is unimodal (it peaks at a single value), symmetrical (symmetrical variations on both sides) and ideally ranging between $\pm\infty$.

  A significance test on the data of each segment can assess the normality of ΔT_1 PDF. These tests generally compare the scores of a sample to a normally distributed set of scores with the same mean and standard deviation, where the null hypothesis (H_0) is "sample distribution is normal".

  One of the most common normality tests is Kolmogorov-Smirnov test (K-S test). By the way, even with the Lilliefors correction, K-S test is not enough accurate as it is very sensitive to extreme values (Ghasemi and Zahediasl, 2012). The Shapiro-Wilk test (S-W test) provides a bigger power than the K-S test (Steinskog et al., 2007). The PDF of a segment that rejects H_0 is very different from the ideal normal distribution and thus it suggests the existence of systematic causes.
- *PDF details.* Once rejected H_0, it is possible to characterize the PDF skewness (lack of symmetry) and kurtosis (pointiness). An asymmetrical PDF with a long tail to the right (higher values) has a positive skewness, while an asymmetrical distribution with a long tail to the left (lower values) has a negative skewness. A flatter distribution has a negative kurtosis, while a distribution more peaked than a normal distribution has a positive kurtosis. According to the target of the method, PDFs of segments with a long tail to the right are critical as they represent waiting time greater than the standard time. At the same time, flatter PDFs represent a big undesired variability.

Table 1 shows a 6-classes structure profile, combining all the possible exits of the test and the PDF details.

2.3. Frequency analysis

Once defined the relative frequency of the segments F_i as the ratio between the operations number of the segments and the total operations, Table 2 shows 4 classes of frequency. The handler's decision maker defines the limits of the classes, to represent the impact that the segment has in terms of the percentage of first baggage to process within a target waiting time, i.e. L_1^*: the larger the frequency F_i, the more critical the class.

2.4. Criticality analysis

The last step consists in combining the results of the two analyses, i.e. structure class and frequency class analysis, in order to obtain a synthetic evaluation. The heat map in Table 3 offers a two dimensional holistic and concise classification to prioritize strategic decisions. The colours represents the segments' contribution to the performance.

3. Case study

The methodology has been applied to a ground handler working

Table 1
Structure profile for the segments PDFs.

Shapiro-Wilk test	Skewness and Kurtosis	Time limit respected	Structure class
Reject H_0	(Skewness > 0) V (Kurtosis < 0)	NO	A
Reject H_0	(Skewness < 0) V (Kurtosis > 0)	NO	B
Accept H_0	-	NO	C
Reject H_0	(Skewness > 0) V (Kurtosis < 0)	YES	D
Reject H_0	(Skewness < 0) V (Kurtosis > 0)	YES	E
Accept H_0	-	YES	F

Table 2
Frequency profile for the segments.

Relative frequency [F_i]	Frequency Class
$F_i \geq (1 - L_1^*)$	A
$\frac{(1 - L_1^*)}{2} \leq F_i < (1 - L_1^*)$	B
$\frac{(1 - L_1^*)}{10} \leq F_i < \frac{(1 - L_1^*)}{2}$	C
$0 < F_i < \frac{(1 - L_1^*)}{10}$	D

Table 3
Heat map for segments criticality assessment.

		Frequency class			
		A	B	C	D
Structure class	A	AA	AB	AC	AD
	B	BA	BB	BC	BD
	C	CA	CB	CC	CD
	D	DA	DB	DC	DD
	E	EA	EB	EC	ED
	F	FA	FB	FC	FD

Criticality class	Characteristics of the class
1	Segments with very high negative contribution to overall performance
2	Segments with high negative contribution to overall performance
3	Segments with moderate negative contribution to overall performance
4	Segments with minor negative or positive contribution to overall performance

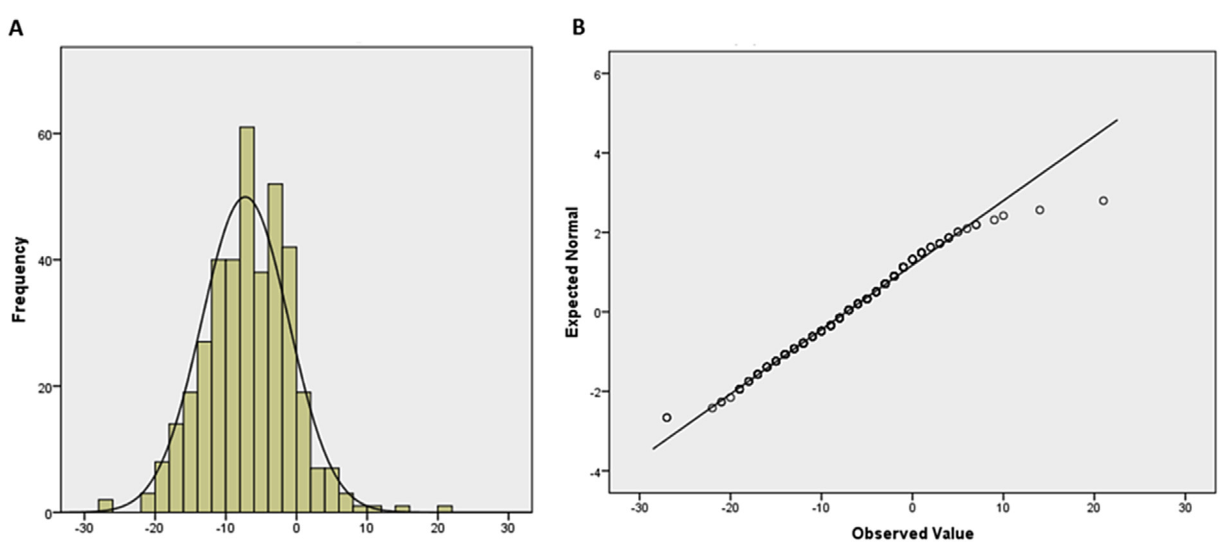

Fig. 1. A) PDF of the whole dataset and normal plot. B) Q-Q plot of the whole dataset.

Table 4
Overall statistics for first baggage delivery time.

	Statistic	Std. error
Mean	−7.25	0.314
Median	−7	
Std. Deviation	6.165	
Minimum	−27	
Maximum	21	
Skewness	0.159	0.124
Kurtosis	1.195	0.248
90th percentile	−0.5	

at Rome FCO (38291989 passengers/year), one of the biggest European airports (ENAC, 2015). The data are anonymous in order not to break the industrial privacy of the parties.

The ground handler serves about 400 flight arrivals per month. More specifically, in the month under observation, it served 386 movements in 2 different terminals, using 13 conveyor belts. The operations refers to 31 landing flights of 9 different airlines, 9 different aircraft, coming from 19 different departure airports.

The airport managing body defined $L_1^* = 90$ (on 100 occasions) and two different targets: $T_{1_{nat}}^* = 19$ minutes for national flights and $T_{1_{intern}}^* = 26$ minutes for international flights (ENAC, 2015). This distinction is not relevant when considering ΔT_1 PDF, which describes the deviations of the process performance from these two values.

Fig. 1A represents ΔT_1 PDF. Table 4 shows how the ground handler respects the limits of the global process in a borderline condition (90th percentile = −0.5). Note that the *p-value = 0.001* of the Shapiro-Wilk test, performed on the whole dataset with a significance value 0.05, rejects H_0 and implies therefore that data are not normally distributed. Fig. 1B confirms this assumption by a Q-Q plot, a quantile-quantile plot to visually check normality (Dhar et al., 2014).

The segmentation of the process helps acquiring an in-depth characterization of the problem.

Table 5 describes the Airport segmentation, which allows defining some criticalities. In detail, Airport 7–9–14 (Criticality Class 3) and especially Airport 12–15 (Criticality class 1) are the most critical ones.

Table 6 describes the Airline segmentation. In this case, it is clear that Airline 5–8 are critical and they operate, as expected, on Airport 9–12–5.

Furthermore, the Flight Number segmentation, i.e. a code to hide the flight ID, offers the more critical results for some of the flights served by Airline 5 (Flight 5–1 and Flight 5–4) and by Airline 8 (Flight 8–2), as shown in Table 7, which serve Airport 7–12–15.

Other segments (e.g. conveyor belt, flight time, date, and day) do not offer any further information and will not be presented.

According to these results, a pool of experts of the ground handler defined some practical guidelines of continuous improvement. These interventions should not be limited to segments that fail to comply with quality standards but also aim at fixing segments with an out of statistical control performance variability.

Table 5
Airport segmentation.

Airport	Relative Frequency	Shapiro-Wilk test2	Skewness	Kurtosis	90th percentile	Criticality Class
Airport 1	0.0155	Accept H0	-0.3707	-10.832	-1	FC
Airport 2	0.0026	Accept H0			-9	FD
Airport 3	0.0803	Accept H0	-0.1202	0.3268	-4	FB
Airport 4	0.0052	Accept H0			-2	FD
Airport 5	0.0699	Accept H0	-0.2217	0.8717	-2	FB
Airport 6	0.0648	Accept H0	-0.5732	-0.37	-2	FB
Airport 7	0.0078	Accept H0	16.301	0.9678	9	CD
Airport 8	0.2435	Accept H0	-0.3148	0.1276	-1	FA
Airport 9	0.013	Accept H0	0.3926	0.0446	4	CC
Airport 10	0.0026	Accept H0			-5	FD
Airport 11	0.0078	Accept H0	-17.321	-0.005	-19	FD
Airport 12	0.2021	Reject H0	87.833	774.240	3	AA
Airport 13	0.0751	Accept H0	0.2456	0.5227	-5	FB
Airport 14	0.0026	Accept H0			7	CD
Airport 15	0.0026	Accept H0			-1	FD
Airport 16	0.1891	Reject H0	12.527	45.170	1	AA
Airport 17	0.0078	Accept H0	0.9352	0.345	-2	FD
Airport 18	0.0052	Accept H0			-1	FD
Airport 19	0.0026	Accept H0			-8	FD

Table 6
Airline segmentation.

Airline	Relative Frequency	Shapiro-Wilk test	Skewness	Kurtosis	90th percentile	Criticality Class
Airline 1	0.1554	Accept H0	0.029	0.2991	-4.5	FA
Airline 2	0.0026	Accept H0			-8	FD
Airline 3	0.0648	Accept H0	-0.5732	-0.37	-2	FB
Airline 4	0.0699	Accept H0	-0.2217	0.8717	-2	FB
Airline 5	0.2591	Reject H0	99.121	988.216	3	AA
Airline 6	0.0026	Accept H0			-4	FD
Airline 7	0.1114	Accept H0	0.516	0.1172	-2	FA
Airline 8	0.3212	Reject H0	0.5359	32.880	-1	DA
Airline 9	0.013	Accept H0	-0.1384	-0.6279	-2	FC

Flight 8–2 and Flight 5–4 illustrate two examples of intervention.

The company started handling Flight 8–2 since three months. This flight is generally served by MD-80 that requires a specific procedure for handling baggage, because of its hold size and structure. This observation addressed the decision maker to organize additional training courses for the operators, which still encounter difficulties during the process.

The actual arrival time of Flight 5–1 usually differs from the

Table 7
Flight Number segmentation for Airline 5 and Airline 8.

Flight Number	Relative Frequency	Shapiro-Wilk test	Skewness	Kurtosis	90th percentile	Criticality Class
Flight 5 - 1	0.0751	Accept H0	-0.8289	0.0546	1	CB
Flight 5 - 2	0.044	Accept H0	0.1112	13.366	3	CC
Flight 5 - 3	0.0026	Accept H0			0	FD
Flight 5 - 4	0.0803	Reject H0	55.571	309.188	4	AB
Flight 5 - 5	0.0026	Accept H0			7	CD
Flight 5 - 6	0.0052	Accept H0			-2	FD
Flight 5 - 7	0.0026	Accept H0			-9	FD
Flight 5 - 8	0.0078	Accept H0	-17.321	-24.530	-19	FD
Flight 5 - 9	0.013	Accept H0	0.3926	0.0446	4	CC
Flight 5 - 10	0.0026	Accept H0			-5	FD
Flight 5 - 11	0.0026	Accept H0			-1	FD
Flight 5 - 12	0.0078	Accept H0	16.301	13.450	9	CD
Flight 5 - 13	0.0026	Accept H0			-3	FD
Flight 5 - 14	0.0104	Accept H0	-0.2287	-38.690	-1	FC
Flight 8 - 1	0.044	Accept H0	-0.6871	20.671	-5	FC
Flight 8 - 2	0.0777	Reject H0	13.921	51.154	1	AB
Flight 8 - 3	0.0751	Accept H0	0.4161	-0.297	-1	FB
Flight 8 - 4	0.0751	Accept H0	-0.6621	0.1969	-3	FB
Flight 8 - 5	0.0466	Accept H0	-0.5275	-0.8558	1	CC
Flight 8 - 6	0.0026	Accept H0			-1	FD

estimated one, often piling up with Flight 5–4. When this happens, the operators' workload increases more than expected, generating delays. This analysis guides the decision maker to ask the Airline 5 for a more precise flight plan and, in the meanwhile, to provide additional resources in that shift.

The method highlighted some important information about the process. However, in order to verify the effectiveness of the improving actions and potential emerging problems, it is necessary to repeat it periodically.

4. Conclusions

Ensuring high quality service levels in all its processes is the top priority for a ground handler, according to the strict requirements that airport managing bodies ask to their operators. Baggage handling, as well as other performance, influences both airlines in terms of delayed aircraft turnaround and passengers satisfaction (Park, 2007), resulting in a poor perceived quality of airport services. Meeting the mandatory standards is a crucial point for the competitiveness and success of the whole industry.

This paper develops a methodology to evaluate in a systematic way the performance variability of handling services with respect to the quality standard. Its application on a real case study shows its benefits to guide decision makers evaluating critical segments of the process and taking proper mitigating actions. The statistical analysis eases the link between research and industrial management, giving a path to identify early signs of criticality, even if in compliance with the standards.

References

Airport Research Center, 2009. Study on the Impact of Directive 96 / 67 / EC on Ground Handling Services 1996–2007.

Bezerra, G.C.L., Gomes, C.F., 2015. The effects of service quality dimensions and passenger characteristics on passenger's overall satisfaction with an airport. J. Air Transp. Manag. 44–45, 77–81. http://dx.doi.org/10.1016/j.jairtraman.2015.03.001.

Burghouwt, G., Poort, J., Ritsema, H., 2014. Lessons learnt from the market for air freight ground handling at Amsterdam Airport Schiphol. J. Air Transp. Manag. 41, 56–63. http://dx.doi.org/10.1016/j.jairtraman.2014.06.016.

Chen, F.Y., Chang, Y.H., 2005. Examining airline service quality from a process perspective. J. Air Transp. Manag. 11, 79–87. http://dx.doi.org/10.1016/j.jairtraman.2004.09.002.

Costantino, F., Di Gravio, G., Tronci, M., 2013. Return on quality: simulating customer retention in a flight firming project. J. Air Transp. Manag. 27, 20–24. http://dx.doi.org/10.1016/j.jairtraman.2012.11.003.

Dhar, S.S., Chakraborty, B., Chaudhuri, P., 2014. Comparison of multivariate distributions using quantile–quantile plots and related tests. Bernoulli 20, 1484–1506. http://dx.doi.org/10.3150/13-BEJ530.

ENAC, 2015. Dati di Traffico 2014.

EU, 2013. P7_TA(2013)0116-Groundhandling Services at European Union Airports.

EU, 1996. Council Directive 96/97 EC.

Ghasemi, A., Zahediasl, S., 2012. Normality tests for statistical analysis: a guide for non-statisticians. Int. J. Endocrinol. Metab. 10, 486–489. http://dx.doi.org/10.5812/ijem.3505.

Park, J.-W., 2007. Passenger perceptions of service quality: Korean and Australian case studies. J. Air Transp. Manag. 13, 238–242. http://dx.doi.org/10.1016/j.jairtraman.2007.04.002.

Soames, T., 1997. Ground handling liberalization. J. Air Transp. Manag. 3, 83–94. http://dx.doi.org/10.1016/S0969-6997(97)00008-2.

Steinskog, D.J., Tjøstheim, D.B., Kvamstø, N.G., 2007. A cautionary note on the use of the Kolmogorov–Smirnov test for normality. Mon. Weather Rev. http://dx.doi.org/10.1175/MWR3326.1.