The Journal of Transport Literature

www.journal-of-transport-literature.org

ITPS
International Transport Planning Society

Peak hour evaluation – a methodology based on Brazilian airports

Dilson Pereira Paulo Filho[1,+]; Carlos Müller[2]; Paulo Tsai Wang[1]

[1] *Infraero, Brasília, Brasil*

[2] *Instituto Tecnológico de Aeronáutica, São José dos Campos, Brasil*

Article Info	Abstract
	This paper aims to establish a new methodology to calculate the design peak-hour passenger based on Brazilian Airport data. First, a cluster analysis using the Ward Hierarchical Method is applied grouping similar airports in terms of annual passenger throughput and EPH (Equivalent Peak-Hour). Then, we proceed with the calculation of the coefficients of variation of the aggregated hourly passenger throughput of the last seven years of the airports in each cluster. We propose the peak-hour for each airport cluster to be determined at the point where the stability of these coefficients is reached. We conclude by estimating a relationship of our proposed design peak-hour passenger as a function of the variables used to determine the clusters.

+ Corresponding author. Instituto Tecnológico de Aeronáutica. Praça Marechal Eduardo Gomes, 50. 12.280-250 - São José dos Campos, SP - Brasil.

E-mail address: dilsonpfilho@gmail.com.

Introduction

The estimation of the design peak-hour passenger is critical for the design of airport terminals and their accessibility as well as for the evaluation of the level-of-service provided by an airport existing infrastructure (Piper, 1990; McKelvey, 1988; Yen, 1995). The concept of the design peak-hour passenger is associated with an hourly passenger throughput although below the absolute peak-hour recorded in a year, but still sufficient to ensure adequate level-of-service at the great majority of the operation time of the airport.

The level-of-service provided to passengers is directly related to the critical moments in the airport operations and how these critical moments are perceived. There are service-standards developed and published, for example, by IATA, BAA and Aéroports de Paris (Ashford, 1988) to be used as benchmark; obviously, in off-peak hours, it is simple to assess the level-of-service, but generally at those moments the quality-of-service is very high, therefore the airport authority should consider to define critical periods of the year to assess the quality-of-service in order to evaluate if such quality is reasonable or if it is necessary to plan for expansions.

This paper establishes a new methodology for determining the design peak-hour passenger for airports based on cluster analysis by Ward Hierarchical Method, by grouping similar airports regarding annual passenger throughput and their EPH (Equivalent Peak Hour), a number derived from the division of the typical day throughput by its highest hourly throughput (Wang, 2012).

The next step is to calculate the coefficient of variation of the hourly passenger throughput during the period from 2005 to 2011 in order to identify at what point stability in these movements is reached. Such level of stability is the key point for the methodology to estimate the design peak-hour passenger by means of a regression using as explanatory variables the same ones that were used to determine the clusters.

1. Literature review

In 1976, ICAO (International Civil Aviation Organization) established a study group (GE / TRAP) to investigate the traffic peak on an international basis and define approaches to improve the situation. In 1978, AACC (Airports Associations Coordinating Council) and IATA (International Air Transport Association) decided to collaborate on a study of peak-hour and airport capacity use, producing a first edition with guidelines for airport management and an updated version in 1990: Guidelines for Airport Capacity / Demand Management.

Peak-hour passenger movement instead of annual passenger movement is the basis for the design of the passenger terminal and its facilities. Furthermore, the operational costs involved in running these facilities are also determined by the peaks. The most important role of operational management of an airport is to maximize the use of existing facilities and minimize the problem of congestion.

Passengers peak-hours should not be considered a problem, in reality it is a phenomenon, i.e., no matter what the political or charging structure is present in an airport, there will always be a tendency to concentrate movements in certain periods of the day due to many factors such as: interest of passengers, airlines planning, fleet usage maximization, etc.

According to Ashford (1997) some factors are crucial to establish the movement of passengers in peak hours: Domestic/International; Traffic Characteristics; Geographical Location; Hub; Catchment Area; and Terminal Capacity.

According to Brunetta (1999) design peak-hour passenger is usually defined from historical data, and it may be the 30th busiest time of the year. Ashford (1997) has outlined below the three most important methodologies for determining design peak-hour passenger: The Standard Busy Rate (SBR), Busy Hour Rate (BHR) and Typical Peak-Hour Passenger (TPHP).

The problem presented by this latter methodology is the discontinuity of the curve. For instance, an airport with an annual throughput of 29,999,999 passengers derives a design peak-hour passenger of 12,000 passengers, while an airport with an annual throughput with one more passenger (30 million/yr) would derive a design peak-hour passenger of 10,500 passengers, a quite different figure for just one more passenger a year.

Wang (1999) developed for Brazilian airports a criterion based on the assumption that highest peak-hours at airports are random at certain levels, i.e. the highest absolute peak-hours in a given year behave in a random way from year to year. To forecast peak-hour demand to a statistically significant level it is necessary to achieve a certain stability. By means of the calculation of the coefficient of variation applied on 48 Brazilian Airports, Wang concluded that stability would be achieved at 96.5% of the annual passenger throughput. So far Wang's Methodology is the only with statistical criteria to choose and to calculate the design peak-hour, therefore this paper uses the same principle to measure the stability of the data.

A great variety of methods is used to determine the design peak-hour at airports, but none of them provides unrestricted service for the peak-hour, as this may result in wasted resources. There is a consensus that planning should aim to meet the demand at some level below this absolute peak, resulting that most passengers receives adequate service levels, and only a small percentage would experience the impact of congestion during very short periods of time.

2. Data Analysis

This section is divided into three subsections: 1) presentation of the variables used, 2) cluster analysis for 56 INFRAERO's airports based on the 2011 passenger annual throughput; and 3) coefficient of variation of passenger movements for each cluster to establish passenger peak-hour.

The preparation of the data leads to the model to calculate the passenger peak-hour using annual passenger throughput and EPH as explanatory variables.

2.1. Description of the database

The data used for this study were obtained from INFRAERO strategic data base, therefore this paper will not present detailed information but only the results of the analysis.

Basically there are two variables derived from this database of 56 INFRAERO's airports: i) 2011 Annual Passenger Throughput and ii) EPH (Equivalent Peak Hour).

As defined by Wang (2012), the EPH is a measure of the infrastructure usage throughout a typical day, which can be used for the evaluation of capacity and of the need for future investments. EPH can be calculated by adding up all the 24 medians (50th percentile) of the passenger hourly throughput of every day of a given year and then dividing it by the maximum value of these medians. High EPH values mean high usage of the infrastructure along the day and/or high homogeneity of the data. On the other hand, low EPH is derived from low usage of the infrastructure and/or high hourly throughput concentration. Low EPH is not desired by infrastructure providers.

2.2. Cluster Analysis

The cluster analysis aims to divide the elements of a group of interest into sub-groups of elements sharing similar characteristics and presenting heterogeneity when compared to elements of another sub-group (e.g., see Al-Sultan and Marrof Khan, 1996; Kaufman Roussweuw and, 1990; Koskosidis and Powell, 1992; Laporte et al. 1989).

Techniques for building the clusters are classified according Hair (2009) in two types: hierarchical and non-hierarchical techniques. The hierarchies are used in exploratory analyzes of the data in order to identify possible clusters and the probable value of the number of groups. As for the use of non-hierarchical techniques, it is necessary that the value of the number of groups already be pre-specified by the researcher.

In this paper, we use the hierarchical method of Ward (1963) which is an agglomerative hierarchical technique. Ward's method assumes that at the beginning of the clustering process there are n groups, since each element is considered to be an isolated conglomeration. At each step of the algorithm, the sample elements are grouped to form new clusters in which all elements share similar characteristics. For example, in the first step, two clusters are formed, with n-1 groups; in the second step another cluster is added with n-2 subgroups left. In this algorithm all clusters are formed after n-1 steps.

At each stage of the clustering algorithm, the two more "similar" clusters are combined to form a new conglomerate. In Ward method cluster similarity is measured by the distance between clusters defined as:

$$d\left(C_i, C_j\right) = \frac{n_i n_j}{n_i + n_j} \left(\bar{X}_{i.} - \bar{X}_{j.}\right)' \left(\bar{X}_{i.} - \bar{X}_{j.}\right), \tag{1}$$

Where ni and nj are the sizes of the clusters C_i e C_j when the grouping process is at stage K and $\bar{X}_{i.}$ e $\bar{X}_{j.}$ are the centroids (vector of the variable means) of C_i and C_j , that is,

$$\bar{X}_{i.} = \begin{pmatrix} \bar{X}_{i1} \\ \bar{X}_{i2} \\ \vdots \\ \bar{X}_{ip} \end{pmatrix} \text{ and } \bar{X}_{j.} = \begin{pmatrix} \bar{X}_{j1} \\ \bar{X}_{j2} \\ \vdots \\ \bar{X}_{jp} \end{pmatrix}, \tag{2}$$

where $\bar{X}_{il} = (\sum_{r=1}^{n_i} X_{irl})/n_i$ is the mean of the variable number "l" of the cluster. At each step of the algorithm, the two clusters which minimize the distance above defined are combined. Only one new cluster may be formed in each step.

Hierarchy Property: At each step, each new conglomerate formed is a grouping of clusters formed in the early stages. If two sample elements are grouped in the same cluster at some stage of the clustering process, they will remain grouped in all subsequent stages, that is, once united these elements cannot be separated. Due to the hierarchy property, it is possible to construct a graph called a dendogram, which is the "tree" or the history of the grouping.

The final choice of the number of groups "g" in which the data set should be allocated is subjective. There are some statistical measures that may be used to assist in determining "g". The criteria adopted in this work are: the distance behavior analysis and the similarity level, detailed below:

Distance: In each step of the clustering algorithm, compute the Euclidean distance between the centroids of the clusters that are being formed. As it progresses in the algorithm, the distance between the centroids is increased, which means that the combined groups become less similar. Thus, if one makes a graph of the distances at every stage of the process, it would be able to verify "jump points" relatively large compared to other distance values. These points indicate the ideal time to stop the algorithm, i.e., the number of clusters g and the final composition of the groups.

Similarity: This test is analogous to the previous one, but this time the behavior of the similarity level is monitored instead of observing the distance at each stage. If the groups C_i and C_j are united at a certain stage, the level of similarity between them is defined by:

$$S_{ij} = 1 - \frac{d_{ij}}{\max(d_{kl}, k, l = 1, \ldots, n)}, \tag{3}$$

where $\max(d_{kl}, k, l = 1, \ldots, n)$ is the largest distance between the sample elements in the distance matrix of the clustering process first stage. In this case, the key is to detect points at which there is a sharp decrease in the similarity of the clusters, indicating that the algorithm should be stopped.

One way to assess whether the partition achieves satisfactory requirements of cohesion (or similarity) and isolation (or separation) of the formed clusters is to calculate the sums of total squares, between and within groups, defined by:

Sum of Total Square (SQTotal):

$$\text{SQTotal} = \sum_{i=1}^{g} \sum_{r=1}^{n_i} (X_{ir.} - \bar{X}.)'(X_{ir.} - \bar{X}.), \tag{4}$$

where $X_{ir.}$ is the vector of p measurements observed for the element number "r" of the sample group "i", i.e.

$$X_{ir.} = \begin{pmatrix} X_{ir1} \\ X_{ir2} \\ \vdots \\ X_{irp} \end{pmatrix}, \tag{5}$$

and $\bar{X}.$ is the vector of global means, regardless any partition, i.e.

$$\bar{X}. = \begin{pmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \vdots \\ \bar{X}_p \end{pmatrix}, \tag{6}$$

where $\bar{X}_l = (\sum_{r=1}^{n} X_{rl})/n$ is global average of the variable number, "l" $= 1, \cdots, p$.

The Sum of Total Squares within the partition groups (Sum of the Residual Squares):

$$\text{SQResidual} = \sum_{i=1}^{g} \text{SResidual}_i = \sum_{i=1}^{g} \sum_{r=1}^{n_i} (X_{ir.} - \bar{X}_{i.})'(X_{ir.} - \bar{X}_{i.})'. \tag{7}$$

Total Sum of Squares between the "g" partition groups:

$$\text{SQbetween} = \sum_{i=1}^{g} n_i (\bar{X}_{i.} - \bar{X}.)'(\bar{X}_{i.} - \bar{X}.)'. \tag{8}$$

The Sum of Total Squares is the sum of the residual squares sum with the sum of squares between groups, yielding the total variation. If a good partition is performed, it is expected that the formed groups have internal cohesion but are heterogeneous in comparison to the others. Thus, it is expected that variations within groups (sum of squares within each group, $SResidual_i$, and altogether, $SQResidual$) are small in relation to the total sum of squares, or, equivalently, it is expected that the variation between groups (Sum of squares between groups) represents the majority of the data variation.

2.3. Coefficient of variation - Stability in peak-hour passenger

Considering the clusters established in the previous section, for each cluster the data used will be the passenger hourly throughput for the last seven years, i.e., from 2005 to 2011.

Using the methodologies of determining peak-hour cited in the literature review, it was found that for the 56 airports in this study, no airport peak-hour was below the hundredth highest hourly throughput. Thus, as a safety margin to achieve a stabilization of the passenger throughput, as well as to save processing time, this study will only take into account the 300 highest hourly throughput for each year to compound the peak-hour database. For each ranked hourly throughput there will be performed a calculation of their coefficient of variation which is the standard deviation divided by the mean. By definition, the stability is achieved when this coefficient does not vary more than 0.01 unities for the entire sample, and thus, the hourly throughput (which the stability was achieved) is to be considered as the representative for the respective cluster. The maximum variation of the coefficient was chosen as 0.01 unities, because if this variation is less than 0.01 unities, in the last two clusters the stability would not be able to be achieved for less than 300 hours.

3. Groupings

According to the method of Ward (1963), let us now consider as the criterion of agglomeration the variables: annual passenger throughput and EPH. In the search for a small number of groups, we will analyze the measures of Similarity and Distance for the last 15 steps. Table 1 shows the values of these measures and their percentage change from the previous step to the current. From Table 1, observing the behavior of distances or equivalently the similarities in the various steps of grouping, one realizes that there is a large loss of similarity (44.32% - it is a figure of two digits contrasting previous one-digit-figures from 41st step to 49th step) and a sharp increase in distance or fusion level (116.40%) from step 49 to 50.

Thus, the value of g for the final partition should be greater than or equal to 7, which has a good degree of similarity (72.43%). Considering then 7 clusters, the airports are grouped as follows and are presented in descending order according to the variable: passenger centroids: Cluster 1: SBGR;   Cluster 2: SBBR, SBGL and SBSP;   Cluster 3: SBCF, SBCT, SBFZ, SBKP, SBPA, SBRF, SBRJ and SBSV;   Cluster 4: SBBE, SBBH, SBCG, SBCY, SBEG, SBFL, SBGO and SBVT;   Cluster 5: SBFI, SBLO, SBMO, SBNF, SBNT, SBPV, SBSL, SBTE and SBUL;   Cluster 6: SBAR, SBBV, SBHT, SBIL, SBIZ, SBJP, SBJV, SBMA, SBMK, SBMQ, SBPJ, SBPL, SBRB, SBSN and SBUR;   Cluster 7: SBCJ, SBCM, SBCP, SBCR, SBCZ, SBJU, SBKG, SBPK, SBSJ, SBTF, SBTT and SBUG. To analyze the groups formed, the Total Sum of Squares between and within groups are:

Table 1 - Historical grouping - Method of Ward - Last 15 steps - Standardized Variables Annual Passengers and EPH:

Step	Groups (g)	Similarity	Similarity Var. (%)	Distance	DistanceVar.(%)
41	15	92.49	-0.64	0.50	8.67
42	14	92.38	-0.12	0.51	1.49
43	13	92.32	-0.06	0.51	0.77
44	12	91.51	-0.87	0.57	10.49
45	11	88.50	-3.29	0.77	35.48
46	10	84.94	-4.02	1.00	30.97
47	9	77.33	-8.95	1.51	50.49
48	8	76.74	-0.77	1.55	2.63
49	7	72.43	-5.62	1.84	18.53
50	6	40.33	-44.32	3.98	116.40
51	5	38.48	-4.58	4.10	3.09
52	4	7.39	-80.81	6.18	50.55
53	3	-43.69	-691.54	9.58	55.15
54	2	-63.79	45.99	10.92	13.98
55	1	-433.60	579.76	35.59	225.79

Table 2 - Sum of the squares within each cluster – Ward Method – Variables: Passengers and EPH.

Cluster i	1	2	3	4	5	6	7
SResidual i	0.00	0.18	0.64	0.86	0.29	0.44	0.59

Sum of Residual Squares: $SQResidual = \sum_{i=1}^{7} SResidual_i = 3.00$

Sum of Squares between partition groups: $SQEntre = SQTotal - SQResidual = 107.00$

Dividing the Sum of Squares between groups by the Total Sum of Squares, it is possible to notice that 97.27% of the total variation is explained by the variability of partition, which indicates that we made a good division of the groups, since there were not too many groups and yet we were able to explain more than 95% of the total variation.

Another way to check the heterogeneity of the groups that were formed is to analyze the centroids (means) of the variables in study for each group (Table 3 below).

Table 3 - Clusters centroids and general centroid – Ward Method - Variables Passengers and EPH.

Cluster i	1	2	3	4	5	6	7
Passengers	30,003,428	15,702,673	7,605,945	2,497,769	1,418,528	462,639	93,956
EPH	17.0	13.1	13.6	11.1	7.6	5.4	2.4

General Centroid: Passengers = 3,192,410.1 and EPH = 7,7

From Table 3, it can be noticed that standardization is necessary to transform the variables to compatible scales, since, in numerically terms, "Passengers" are million times "EPH". After standardization both variables were similarly important for the definition of the groups. There are clear differences between the mean values for both groups of passengers as for EPH, while groups 2 and 3 have close values of EPH.

4. Stability in passenger peak hour.

In possession of the coefficients, it seems that the peak hour for Cluster 1 is obtained in the 6th highest hourly throughput, corresponding to 99.88% of annual passenger movement, for Cluster 2 it is the 8th highest hourly throughput, corresponding to 99, 06% of annual passenger movement, for Cluster 3 it is the 14th highest hourly throughput, corresponding to 99.47% of annual passenger movement, for Cluster 4 it is the the 25th highest hourly throughput, corresponding to 99.04% movement of passengers annually; for Cluster 5 it is the 30th highest hourly throughput, corresponding to 97.94% of annual passenger movement, for Cluster 6 it is the 37th highest hourly throughput, corresponding to 96.42% of the movement passengers annually, and for Cluster 7 it is the 57th highest hourly throughput, corresponding to 88.57% of annual passenger movement.

From the above it can be inferred that additional clusters would implicate in higher difficulty of achieving enough stability. In other words, a theoretical cluster 8 in this case would implicate in something around the 100th hour to reach stability, proving mathematically that enough levels of stability are not achievable for every airport.

5. Least Square Regression.

According to the variables presented in this study: annual movement of passengers (PAX) and Equivalent Peak Hour (EPH), the Equation (10) yielded is:

$$hp = \beta 1 pax + \beta 2 eph + \mu \quad (10)$$

Table 4 - Results of the Linear Regression Model.

OLS estimation	Estimates efficient for homoskedasticity only		Statistics consistent for homoskedasticity only	
Number of obs = 56	F(2, 53) = 748.01		Prob > F = 0.0000	
Total (centered) SS=107720515.4	Centered R2 = 0.9658	Total (uncentered) SS = 175159696		Uncentered R2 = 0.9790
Residual SS = 3685689.628	Root MSE = 256.5			
hp1 | Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]
pax | .0002258	9.37e-06	24.10	0.000	.0002074 .0002442
eph | 45.13556	11.90984	3.79	0.000	21.7927 68.47842
_cons | 28.7403	80.39928	0.36	0.721	-128.8394 186.32

It is possible to verify that the variables PAX and EPH are statistically significant at 1%, i.e., there is strong evidence of the causal effect of these variables on passenger peak-hour. Regarding elasticities (P-value ranges: p<0.01 and adj. R-sq : 0.964), it is possible to conclude that annual passenger throughput (Pax) impacts more than EPH, i.e., for every additional 1% of PAX there is an increment of 0,6368% (t-test 0,034) of the peak hour demand, while for every additional 1% in the EPH results in 0,3170% (t-test 0.084) peak hour demand increment.

Conclusion

The methodology proposed by this paper is very useful for defining, mathematically, a design peak hour. A clustering process was chosen because, according to Hair (2009), the best advantages of the hierarchical method is its simplicity, the development of measures of similarity and the results are achieved in little time. The clustering method by Ward, which grouped the airports into seven clusters, presented very significant variables: the annual passenger throughput and the Equivalent Peak Hour (EPH). Other variables were tested such as: declared airport capacity, connecting passenger throughput and international passenger throughput; but the two that were chosen for this study (annual passenger throughput and EPH) were the ones who achieved best results. Another interesting finding was the use of the coefficient of variation to determine the stability of the peak-hour demand for the seven clusters. The independent variables were chosen to explain objectively the phenomenon of airport passenger concentration, other important variables, such as the level-of-service perceived by the users, is very subjective and cannot be mathematically modeled at the statistical level of significance this study required.

After defining the independent variables, it was possible to model a linear regression which it was proven to be, statistically, very significant. The contribution intended with this model is to have a robust and consistent methodology, taking into account the impacts on peak hour demand by the annual passenger throughput and the EPH, two easy variables to get, either in the present time, or by means of econometric scenarios for the future. All of the models from the literature review do not take into account the impact of restrictions derived from airport capacity, this present methodology is the first one to evaluate it and to present a simple model considering EPH as a very good airport capacity proxy.

References

Al-Sultan, K.S. and Maroof Khan, M. (1996). Computational experience on four algorithms for the hard clustering problem. Pattern Recognition Letters 17 (3), 295–308.

Ashford, N.; Stanton, M., H.P. and Moore, C.A. (1997). Airport Operations. Second Edition. Copyright © 2007 Mc Graw Hill.

Ashford, N., (1988). Level of service design concept for airport passenger terminals - a European view. Transportation Planning and Technology 12, 5-21.

Brunetta, L.; Righi, L. and Giovanni, A. (1999). An operations research model for the evaluation of an airport terminal: SLAM (simple landside aggregate model). Journal of Air Transport Management 5, 161–175.

Hair, Joseph F. (2009). Multivariate data analysis. Bookman Compania Editora, São Paulo, SP, Brasil, 6th Edition.

Kaufman, L. and Roussweuw, P. (1990). Finding Groups in Data: An Introductory to Cluster Analysis. Wiley, New York.

Koskosidis, Y. and Powell, W.B. (1992). Clustering algorithms for consolidation of customer orders into vehicle shipments. Transportation Research Part B 26 (5), 365–379.

Laporte, G.; Chapleau, S.; Landry, P.-E. and Mercure, H. (1989). An algorithm for the design of mailbox collection routes in urban areas. Transportation Research Part B 23 (4), 271–280.

Mckelvey, F., (1988). Use of analytical queuing model for airport terminal design. Transportation Research Record 1199.

Piper, D.H.P. (1990). Estimation the passenger flow in departure halls. Airport Forum, May 262–270.

Wang, P.T. and Pitfield, D.E. (1999). The derivation and analysis of the passenger peak hour: an empirical application to Brazil. Journal of Air Transport Management. vol. 5, 135-141.

Wang, P.T. (2012). Peak hour equivalency applied to current Brazilian methods to manage existing infrastructure and to plan future airport expansions. Presented at 2012 ATRS.

Ward, Joe H. (1963). Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association. Vol. 58. No. 301, 236-244.

Yen, J.R. (1995). A new approach to measure the level of service of procedures in the airport landside. Transportation Planning Journal 24 (3), 323–336.