

Research Article

An Application of Cluster Analysis Method to Determine Vietnam Airlines' Ground Handling Service Quality Benchmarks

Tien-Chin Wang¹ and Yen Thi Hong Pham ^{1,2}

¹Department of International Business, National Kaohsiung University of Science and Technology, No. 415, Jiangong Rd., Sanmin Dist., Kaohsiung City 807, Taiwan

²Faculty of Business English, Foreign Trade University, 91 Chua Lang St., Dong Da Dist., Hanoi, Vietnam

Correspondence should be addressed to Yen Thi Hong Pham; phamnhhatminhpro555@gmail.com

Received 13 August 2019; Accepted 13 February 2020; Published 30 May 2020

Academic Editor: Eneko Osaba

Copyright © 2020 Tien-Chin Wang and Yen Thi Hong Pham. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper recommends that Vietnam Airlines use a pro-offered model to both evaluate and improve its current service network being operated at international airports. The model includes cluster analysis, ANOVA, and Scheffé post hoc to provide service performance insights and to serve as a complementary corporate benchmark for evaluating service potential and for identifying deficient service areas. By means of this model, the managerial board can designate a potent strategy for ground handling service. Additionally, the given model provides expatriate station managers with a clearer viewpoint of the localized productivity level as performed in relation to other airports concomitant within their own clusters.

1. Introduction

In Vietnam, both the domestic and the international air travel markets have experienced intensifying competition in the recent past few years. Thus, Vietnam Airlines (VNA), the national carrier, now faces both significant and unprecedented operational challenges on a global scale. Within Vietnam's domestic market, apart from fast-growing local rivalry, market saturation remains a regular obstacle VNA faces. As domestic market expansion contracts, the company has had to consolidate in the international marketplace. Regarding such a development, VNA experiences fierce and direct daily competition from major international carriers such as Singapore Airlines, Japan Airlines, American Airlines, and Lufthansa as well as a slew of local low-cost carriers located in Northeast Asia and Southeast Asia.

Under current circumstances, and an awareness of the distinct challenge posed by international expansion, VNA has introduced further promotional incentives to customers such as mileage reward programs or frequent flyer membership programs. However, the benefits from such marketing strategies erode over time because other airlines have

already offered similar promotions in the past. Realizing the limited impact that marketing strategies may produce, the national carrier now focuses on improving overall customer service quality and satisfaction to attract repeat business. This is the shared purpose of this article by recommending the application of cluster analysis method to the airline operations.

Currently, the airline provides online ticketing and reservations, professional ground services, and in-flight guest services. Among these offerings, airport ground services are of considerable importance since they most often directly impact the customer service experience. VNA has signed a service-level agreement (SLA) to act in accordance with numerous ground handling companies at international airports in order to regulate these services. The carrier also attaches a supervisory expatriate station manager to each ground handling company to supervise the on-site service quality level expected.

The SLA defines the mutually agreed upon set of standards and targeted goals used to monitor the general ground handling company's performance. Regular meetings are then organized between VNA and the handling company to

track the agreed-upon level of service quality against assessed performance standards and regular service quality targets.

The exact indicators or criteria of the different SLAs may tend to vary from site-to-site. However, they trend towards the following: (1) punctuality; (2) check-in process; (3) boarding process; (4) staff attitude; (5) customer complaints; (6) baggage mishandling; and (6) personal documents mishandling.

The first 3 columns of Table 1, shown below, provide a rubric of main indicators voiced in the SLA signed between Vietnam Airlines and China Airlines Ltd., the ground handling company located at Kaohsiung's Siaogang International Airport (KHH) in southern Taiwan for the year 2016.

Even though the SLAs may indicate the putative service standards and service quality targets for handling companies to follow, they still do not guarantee that customers' needs are being met. In order to determine whether or not customers are satisfied with airport ground handling services, VNA conducts regular passenger surveys to obtain feedback related to the service quality provided.

In addition, VNA's Market Service Department personnel maintain comprehensive data records of the handling companies' regular performance levels. The data from these sources and from part of the passenger survey are then consolidated and checked against the set of standards and targets already predetermined. From these results, the airline may decide the nature of ground handling services at airports that are up to standard and those which are not.

The last column in Table 1 is an example of how the service quality of the handling company at Kaohsiung's Siaogang International Airport in Taiwan was evaluated for the year 2016.

Presently, the method VNA utilizes to determine the SLA is by making a comparison of the performance results for each of the ground handling companies with its respective service standards or targets preset as part of the SLA. If the performance results are not lower than the targets, then the carrier concludes that the handling company has met the desired performance criteria. Apart from this rather generalized assumption, there is little that can be concluded from the passenger surveys and recorded data.

In this study, the authors will use the same VNA data that were consolidated in 2015 and 2016; however, other tools were used for purposes of analytics. In this regard, cluster analysis, ANOVA, and Scheffé post hoc were used to provide VNA with a more accurate picture of how these handling companies are performing, especially in relation to the others in their group or cluster and then to those in other clusters. It is also argued that these tools can enable the carrier to establish more achievable targets for each group of ground handling companies when followed by benchmarking for future performance standard levels for the next higher group or subgroup.

The primary contribution of the article is in its suggested methodology and in its approach to evaluate ground handling service quality due to insufficient coverage in the literature because of its relatively small role in airline service

quality evaluation. Such an evaluation may be adapted and then expanded conversely to other service industries. The evaluation method provides an insight into service performance, and it can serve as a company's complementary benchmark for evaluating its own services as well as identifying deficient service areas. Moreover, it affords a definite reflection of the industry's perspective on service quality evaluation, and unlike other rather academic articles related to airline service quality, it insures the practicality of the findings and methods for the industry.

The remainder of this paper is organized as follows. Section 2 explains what cluster analysis is, its benefits, and its application. This section also emphasizes service quality along with criteria used to establish benchmarks for the airline service. Section 3 introduces the research methodology applied, and Section 4 describes and then discusses the data analysis results. Finally, in Section 5 some relevant conclusions are made.

2. Literature Review

2.1. Cluster Analysis. Cluster analysis, or clustering is a statistical method used to classify groups [1]. It is performed to discover groups of objects found in data on the basis of some form of proximity measurement defined among them [2, 3]. Those relative group memberships are created with respect to their proximity to one another. However, in hierarchical clustering, data are not separated into a specific number of classes or groups at a single given phase. As an alternative, the classification procedure includes a series of separations running from one group of all individuals to the n group of a single individual. Both types of hierarchical clustering techniques, i.e., agglomerative methods and divisive methods, have been considered to be the optimal phase in which analysis took place. Agglomerative methods are conducted to form a sequence of successive connections of the n individuals into groups, while divisive methods partition the n individuals into a better form of classification [1]. A process to form hierarchical clusters of reciprocally undivided subgroups is repeated until only one group remains [4]. Regarding the result, a dendrogram is produced which provided a convenient visual aid with which a hierarchical sequence of clustering assignments is exhibited. It appears in the form of a simple tree where each node serves to indicate a cluster. Each cluster represents a single data point, and the root node indicates the cluster consisting of the whole dataset.

Ward's method has been widely employed to conduct cluster analyses in prior research (e.g., [5–8]). Everitt et al. [1] emphasized that out of the seven methods of cluster analysis available, the standard agglomerative hierarchical clustering method (i.e., Ward's method) possessed the greatest marked tendency to collect same-size, spherical clusters and was a sensitive method to identify outliers. Hands and Everitt [9] claimed that this method was a more complicated method, but also provided a more accurate one in terms of results and minimized variance between objects. Euclidean distance was recommended for distance measure and most used

TABLE 1: Main indicators for 2016 service delivery standards and results.

#	Main indicators	Target	Results
1	Punctuality—within 15 minutes of scheduled time of departure (STD)/expected time of departure (ETD) (refer only to flight delays attributable to pertinent handling company)	≥99.5%	99.89
	Check-in process		
	(i) Transit passengers handling		
2	(ii) Queuing area in order	≥84/100	83.70
	(iii) Professional staff at check-in		
	(iv) If required, assist passengers in doing airport procedures		
	Boarding process		
3	(i) Clarity and coherency of speech used for boarding	≥84.5/ 100	82.68
	(ii) Lounge-to-aircraft convenience		
4	Manners and professional attitude expressed by staff	≥84/100	83.74
5	Passenger complaints, per 1,000 passengers	≤0.1	0.00
6	Mishandled baggage, per 1,000 passengers	≤0.2	0.20
7	Mishandling occurring in examination of passengers' travel documents, per 10,000 passengers	≤0.3	0.09

Source: adapted from Kaohsiung Service Delivery Standards 2016 and Results.

whenever applying Ward's method. For this study, Ward's method applied with Euclidean distance presented the clearest image of clustering.

2.2. Airport Classification. Classifying airports according to service capabilities is popular. In air transport research, major scholars have grouped airports according to their common attributes (e.g., [6, 7, 10–14]). They have all made various aims and approaches for classifying airports depending on the purpose of the survey. Classification schemes frequently used groupings in terms of benchmarking practice [10, 14–16].

Although airport classification has already been studied in a variety of ways regarding diverse aspects, current research related to airports seemingly ignored the intangible issue of passenger service quality. Airports have been classified mostly in terms of their connectivity, geographic location, functionality, traffic distribution, airport size, cargo capacity, utilization and technical characteristics, ownership, efficiency and productivity, and network position [6–8, 10–12, 14, 17]. However, throughout the classification process of airports, variables related to passengers have the primary inclusion. For instance, Adikariwattage et al. [10] used the U.S. Bureau of Transportation Statistics survey database with the additional variables of gate numbers, annual volume including passengers of origin-destination, and transfers from both international and domestic to U.S. airports. Furthermore, by considering an airport's position within the network, Malighetti et al. [6] found the presence of strategic groups in their cluster analysis made up of 467 European airports. These researchers used variables related to passenger connectivity, such as the seat availability on scheduled flights, the amount of destinations offered, and traffic distribution amongst routes. Rodríguez-Déniz et al. [13] classified airports based on the available air ticket information taken from more than 30 major U.S. air carriers. Apart from concentrating on passenger variables that were related to airports, some studies have focused on overall airport efficiency. For example, Rodríguez-Déniz and Voltes-Dorta [16]

used output and input pricing, with cost elasticity and factor shares serving as optimal variable weights in order to estimate an airport's efficiency. They identified 17 distinct airport clusters using this given technique. Sarkis and Talluri [14] measured the operational efficiency of 44 major U.S. airports. In this case, four input measures (e.g., airport operational costs, number of airport employees, number of gates, and number of runways) and 5 output measures (e.g., operational revenue, passenger flow, commercial and general aviation movement, and total cargo) were used to provide efficiency measures.

While cluster analysis is a useful tool to group airports, hierarchical clustering is also commonly used in the classification of airports, as employed by Malighetti et al. [6]; Mayer [7]; Rodríguez-Déniz et al. [13]; Sarkis and Talluri [14]; and Vogel and Graham [8]. One of the key benefits of hierarchical clustering over k-means clustering, as remarked by Rodríguez-Déniz et al. [13], is that by illustrating by means of a tree structure, hierarchical classification can indicate a typical structure which is more informative than the flat clusters gained from other splitting methods, such as k-means.

After grouping airports, the usage of airport clusters in benchmarking exercises is a main purpose. Sarkis and Talluri [14] treated the efficiency scores of major U.S. airports across 5 years by way of a clustering method when identifying benchmarks used to improve poorly performing airports. Mayer [7] based his airport cluster analysis upon cargo tonnage throughput. The objective of this study was to provide an insight into the heterogeneity of cargo airports and then to define comparator airports in the air cargo marketplace.

This research is an application of cluster analysis towards the classification of VNA service-linked international airports. This airport hierarchical clustering is based on ground handling service determinants, which was self-determined by VNA. The quality service provided by the ground handling service companies at VNA's international airport destinations represented performance efficiency as evaluated by the Market Service Department's review of VNA passenger surveys and records.

2.3. Service Quality in the Airline Industry. It has been realized that delivering superior service quality can be of key importance to success and survival in today's hypercompetitive business environment. Hill et al. [18] highlighted the impact of service quality on a company's service strategy formulated to improve profitability: "(...) for services, "Production" typically takes place while the service is delivered to the customer. The production function can also perform its activities in a way that is consistent with high product quality, which leads to differentiation (and higher value) and lower costs." (p.92). Moreover, service quality was found to be an independent and positive direct effects on satisfaction [19]. By offering superior service, companies were able to charge premium prices for their product [18, 20] while gaining growth of market share [21]. It was also noticed that companies could still make a profit while reducing the customer defection rate at the same time [18]. Thus, airlines can agree with the motto "Improving service quality is improving profitability."

In reviewing lessons gleaned from past service quality studies, there was an indication that any service improvement whatsoever can increase the customer base through new and repeated purchases from more loyal customers. For this approach to work, Johnson et al. [22] claimed that the prediction of future retention behavior and profitability could be made through a determination of a cumulative customer satisfaction level. In the airline industry, carriers have made considerable efforts to improve overall service quality with a view to satisfying passengers and meeting customer expectations that will serve to maximize long-term profitability. Therefore, service quality needs to be perceptible and assessable by customers (e.g., [23–27]). Gilbert and Wong [24] argued that when flying with an airline, all passengers should share the same positive expectation of desirable service quality. However, this same expectation does not only exist when passengers are not homogeneous in terms of their racial identities and travel objectives. Chen and Chang [23] investigated airline service quality from a process viewpoint by studying the distances between passengers' service expectations and the real service rendered at ground level and from in-flight services. Furthermore, these researchers discovered the distances associated with passenger service expectations and perceptions of these expectations by inquiry of the relevant frontline managers and employees. They stated that there was an existing distance, and fliers seemed to be most concerned about the aspects of responsiveness and assurance available when interacting with the airline's frontline staff. Wu and Cheng [28] designed a purpose-built model to represent passengers' overall perceptions of airline service quality in the industry to provide the most satisfactory experience for passengers as a whole. Apart from the satisfaction factor, passenger expectations and perceptions were examined in relationship to airline service in different contexts, including airline service quality perceived by passengers in an uncertain environment [27]; consumer expectations and perceptions in an international setting [26]; passengers' perceptions of service quality leading to the choice of carriers for international air travel in Taiwan [29]; and passenger expectations and airline services [24]. For the airline service

industry, understanding their passengers' expectations and perceptions necessary to satisfy them is of key importance because this is precisely what keeps the passengers repeatedly flying with the airlines, which inevitably leads to profitability.

2.4. Criteria and Subcriteria for an Airline Service Quality Evaluation. The literature has made considerable progress as to how service quality perceptions are to be measured (e.g., [25, 30–36]). Parasuraman et al. [25] used terms such as reliability, responsiveness, empathy, assurances, and tangibles to best describe service encounter characteristics. These 5 criteria are well known under the name of the SERVQUAL model and integrate 22 subcriteria used to measure functional service quality levels. This model has been widely used in the airline service (e.g., [24, 37–39]), and it was followed by a number of other models constructed over time, such as with the performance-based SERVPERF model developed by Cronin and Taylor [31] or SERVPEX which can measure disconfirmation in a single model [32].

The business of various airlines depends on the quality of the services they afford. An evaluation of such service quality may be formulated using multiple criteria, and multiple subcriteria, to cover the entire business activities of an airline (e.g., [24, 37–39]). Gilbert and Wong [24] conducted a study on passengers' service expectations while in the Hong Kong airport and used 7 dimensions, with 26 scaled questions, as a management diagnostic tool of its service quality. Gupta [38] carried out research with 29 subdimensions under further 7 main dimensions, which represented the largest number of dimensions and subdimensions ever used to evaluate Indian airline industry. Chou et al. [37] tested their model through a Taiwan airline at the Siaogang International Airport in Kaohsiung, Taiwan. They evaluated the airline service quality, using 28 items under 5 major dimensions. However, Tsaor et al. [39] pointed out that the real meaning of quality in airline services was difficult to describe and to measure because of its heterogeneity, intangibility, and inseparability.

According to the airline industry, the major criteria used to evaluate service delivery systems contain management, staffing, passenger satisfaction, reliability, and tangibility which are all found in previous research [38–41]. These 5 criteria covered the whole airline business including ground service and in-flight service. Ground handling was considered as a part of ground service and took place in an airport context. Criteria and subcriteria used to measure ground handling service provided to passengers at airports were relatively consistent in nature, and they have been frequently employed in many empirical studies. Research commonly involved on-time performance, the check-in process, the boarding process, staffing attitude, customer complaints, and baggage mishandling incidents [23, 24, 28, 37–45]. Zhang et al. [46] and Bowen and Headley [42] focused on 5 determinants which included on-time arrivals, mishandled baggage (flight delays), involuntarily denied boarding (oversales), and consumer complaints. Bowen and Headley [42] even used 12 subelements for the determinant named as consumer complaints. However, the findings of both studies were not consistent because different weightings were used

in Bowen and Headley's research. Moreover, staffing attitude was a key aspect to measuring the appropriate level of staff-member interaction with passengers. Positive feelings and negative feelings towards staff behavior played an important role in determining customer satisfaction [43, 44]. Likewise, on-time performance was of similar importance. An airline's achieved punctuality was a major source of passenger satisfaction concerning airline's perceived reliability [47, 48]. Additionally, Wu and Cheng [28] stated that passengers generally considered their eventual waiting time as a vital criterion in evaluating their service quality experience. Furthermore, personal interactions between passengers and airline employees will likely impact positively or negatively on passengers' perceptions of the airline service quality. Communication attributes used in much of the research to describe regular interactions between airline staff and passengers like the professional appearance of the staff [40], staff with professional knowledge skills, [39, 43, 49], staff with consistent service, willingness to help [29, 50], and language ability [29] were considered to be extremely sensitive to fliers.

The aforementioned empirical research was related to airlines operation, which may include ground handling services. However, there were no studies which had covered ground handling activity on a sole basis, to which this paper strove to investigate.

3. Methodology

3.1. Research Design. The proposed model is comprised of cluster analysis taken along with Ward's method and Euclidean distance, ANOVA, and Scheffé post hoc test. Hence, the research design was divided into 5 steps (see Figure 1).

Step 1: this step is to select criteria for analysis. In this study, 6 criteria were adopted.

Step 2: for airport selection in this step, VNA's 28 international airports were chosen due to their sufficient data values (see Table 2).

Step 3: cluster analysis was used to group all of the airports according to the given evaluation criteria. Ward's method has been trialed with different distances, but it was only with Euclidean distance that the dendrograms presented the clearest tree-like graphics of clustering.

Step 4: ANOVA was applied in order to display criteria defined by a significant performance gap. Hence, an overview of the service was presented.

Step 5: the Scheffé post hoc test was applied to examine the criteria with significant differences in order to determine the cluster performance levels. Based on these levels, the benchmarks and possible targets for next period could be readily defined.

3.2. Sample Characteristics. Research data have been collected from VNA's Market Service Department. It is comprised in large part of quarterly SLA service quality reports for

ground handling companies at VNA's 28 international airport destinations. From these quarterly figures, the authors formed two sets of data by average, one for the year 2015 and the other for 2016.

The majority of sample airports are located in Asia, which accounts for 80% (23 airports) of airports concerned in the study, while the balance of 14% (4 airports) are located in Europe, with a further 6% (2 airports) located in Australia.

The VNA Market Service Department has developed 7 criteria to assess the service quality provided by ground handlers. Based on these criteria, the information used to form the data rooted from different sources included the information system used both internationally and internally, daily reports produced by station managers, and autogenerated messages relevant to routine operational activities. The criteria used are described as follows:

- (i) Check-in process, boarding process, and staffing attitude: the figures for the first 3 criteria originate from completed passenger surveys administered quarterly and distributed randomly onboard VNA flights. In the survey, passengers are requested to report the amount of time spent waiting at check-in counters and standing in queue for their check-in turn; the perceived level of convenience concerning boarding time; the gate-area staff's assistance by providing instructions coupled with information signage and multi-lingual announcements; and the perceived helpfulness and the courteous nature of airport staff during their communications and interactions with passengers.
- (ii) Customer complaints: data with respect to this criterion were derived from internal record keeping, indicated by the rate of consumer complaints per 1,000 passengers.
- (iii) Baggage mishandling: the figures for this criterion were drawn from the *World Tracer* system, which traces lost baggage items worldwide. This criterion indicates the rate of mishandled baggage reported per 1,000 passengers.
- (iv) Personal documents mishandling: the data related to this criterion is also drawn from VNA's internal records. This criterion indicates errors which occur in the checking of passengers' travel documents in compliance with the entry, in transit, or with other requirements based on countries-of-origin/arrival.
- (v) Punctuality: the actual time an airplane departed, or arrived, as compared to the scheduled time of departure (STD) and expected time of departure (ETD) that define if the flight was on time or not. A flight is considered as late if the time difference is fifteen minutes or more. The actual time is measured by the automatic messages sent system wide when an airplane takes off or lands.

Although there are 7 criteria that have been commonly used by the literature as previously-stated, only the first 6 criteria were studied in this study. The criterion related to

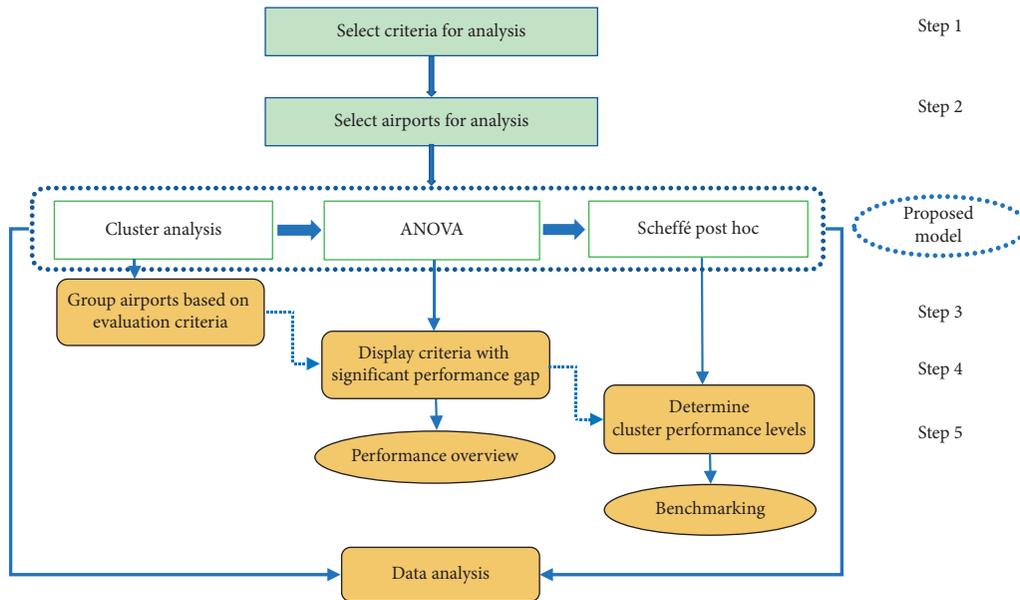


FIGURE 1: Proposed model of research design.

TABLE 2: Airport names according to IATA codes.

IATA code	Airport name
BKK	Suvarnabhumi International Airport (Bangkok)
CAN	Guangzhou Baiyun International Airport
CGK	Chongqing Jiangbei International Airport
CTU	Chengdu Shuangliu International Airport
DME	Domodedovo International Airport (Moscow)
FRA	Flughafen Frankfurt
FUK	Fukuoka Airport
HKG	Hong Kong International Airport
HND	Tokyo Haneda Airport
ICN	Incheon International Airport (Seoul)
KHH	Kaohsiung International Airport
KIX	Kansai International Airport (Osaka)
KUL	Kuala Lumpur International Airport
LHR	Heathrow Airport (London)
LPQ	Luang Prabang International Airport
MEL	Melbourne Airport
NGO	Chubu Centrair International Airport (Nagoya)
NRT	Narita International Airport (Tokyo)
PEK	Beijing Capital International Airport
PNH	Phnom Penh International Airport
PUS	Gimhae International Airport (Busan)
PVG	Shanghai Pudong International Airport
REP	Siem Reap-Angkor International Airport
RGN	Yangon International Airport
SIN	Singapore Changi Airport
SYD	Sydney Airport
TPE	Taiwan Taoyuan International Airport
VTE	Wattay International Airport (Vientiane)

“Punctuality” is excluded because almost all airports were on time, to some greater or lesser extent.

The figures for these criteria are calculated in terms of a 100-point score scale. For the first 3 criteria, the scores are shown to represent the effectiveness of the ground handling companies, whereas in the last 3 criteria, the scores represent their perceived ineffectiveness.

By comparison with the company targets through both years, the data values indicate that the system reached only a median number of the preset targets. The achievement of targets outnumbered in the last 3 criteria, especially the rate of customers’ complaints was considerably low.

4. Research Findings and Discussion

4.1. Airport Classification Based on 2015 and 2016 SLA Service Quality. As previously mentioned, cluster analysis using Ward’s method as the applied agglomerative algorithm and Euclidean distance as distance measurement is selected for this analysis. Although both figures (see Figure 2 for 2015 and see Figure 3 for 2016) revealed cluster analysis results with numerous clusters at different distances, the authors abbreviated a cutoff on the dendrograms at the distance of 12 for both years, respectively. In this way, there were always 5 airport clusters to be observed for each year. The distance was chosen to benefit operations management, and it was out of consideration for the visual presentations that the airport graphics could be quickly joined to form 5 main groups. This method showed a higher degree of homogeneity of the airports in each cluster in terms of efficiency. The 5 resulting clusters in 2015 were named A1, A2, A3, A4, and A5, and another 5 in 2016 were B1, B2, B3, B4, and B5. According to their performance, as measured by VNA, the features of these clusters are described as follows:

Cluster A1: RGN, PNH, KUL, and HKG (4 airports)

The cluster’s performance needed some improvement in the first three attributes, including check-in, and boarding, and staff attitude. It was on average with baggage handling and personal travel documents. The number of cases filed by passengers was lower than

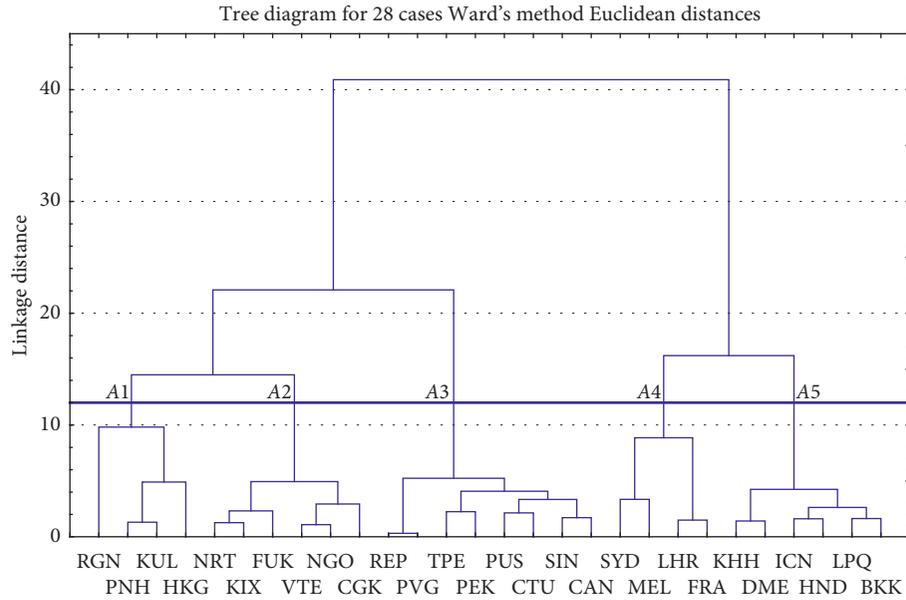


FIGURE 2: SLA performance in 2015.

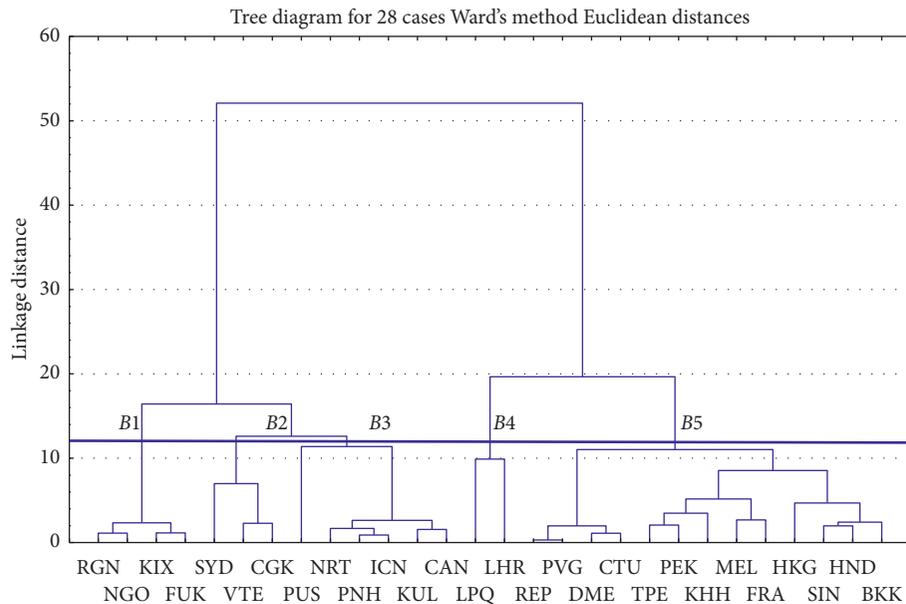


FIGURE 3: SLA performance in 2016.

expected. Less than 40% of the VNA targets set for this cluster were reached.

Cluster A2: NRT, KIX, FUK, VTE, NGO, and CGK (6 airports)

The cluster exceeded performance standards in baggage handling. According to the other criteria, airport performance was in median level. Moreover, performance met the airline’s expectations in terms of functionality based on customer complaints and mishandled baggage attributes. It achieved over 40% of airline’s targets.

Cluster A3: REP, PVG, TPE, PEK, PUS, CTU, SIN, and CAN (8 airports)

The cluster was above performance standards in carrying out tasks connected to the boarding process, yet remaining relatively unproductive in terms of check-in handling and staffing attitude. The majority of airports in this grouping received no complaints from passengers. It could possibly reach 40% of the airline’s targets.

Cluster A4: SYD, MEL, LHR, and FRA (4 airports)

The cluster met 50% of the targets set by the airline. It was among the first place finishers in the system in terms of carrying out tasks related to passenger check-in, the boarding process, and staffing attitude. However, the rates for mishandled baggage and personal documents were considered to be highest.

Cluster A5: KHH, DME, ICN, HND, LPQ, and BKK (6 airports)

The cluster provided above average service quality. Airports in this grouping achieved over 65% of company targets. Like A4, the cluster gave performance exceeding standards for the first three criteria. It also had above average performance levels in those criteria remaining.

Cluster B1: FUK, KIX, NGO, and RGN (4 airports)

For the check-in, the boarding process, and staffing courtesy, B1 needed improvement because it was in the lowest ranking of the system. Its performance delivery according to the last 3 attributes was better than the first three groupings. In addition, the cluster reached less than 40% of the stipulated targets.

Cluster B2: CGK, SYD, and VTE (3 airports)

Although the cluster achieved most of targets envisioned (65%), this grouping's boarding process control needed improvement. The performance levels for check-in, the amount of poor comments, and the helpfulness of staffing were all above average.

Cluster B3: CAN, ICN, KUL, NRT, PNH, and PUS (6 airports)

In addition, this cluster reached 30% of targets; this means that there was average performance regarding most criteria in this cluster.

Cluster B4: LPQ and LHR (2 airports)

This cluster is comprised of two high level airports reporting the handling of check-in service, the boarding process, and staff communicating with passengers. It could reach over 50% of the targets laid out. It should be noted that one airport was much better in carrying out baggage handling and personal travel document control than the other.

Cluster B5: REP, PVG, DME, CTU, TPE, PEK, KHH, MEL, FRA, HKG, SIN, HND, and BKK (13 airports)

This cluster had productive performance evaluation in the first four criteria. No passenger complaints were made against this cluster, with the possible exception of BKK. In addition, most airports located in this cluster were able to consistently provide a high level of baggage handling and personal travel document handling. This grouping reached less than 50% of the set targets.

In conclusion, the amount of the both-year-targets hit was underexpected. More importantly, the performance of the system was demonstrated in the way that various performance levels exist among 5 of the 2015 clusters and 5 of the 2016 clusters for individual criteria, while each cluster showed similar achievement levels of performance in one criterion, or some criteria, shared among airports.

4.2. Specification of Service Quality Levels. After the 10 clusters were produced, ANOVA has been applied to discover in which criteria there existed statistically significant differences of service quality. Then, for criteria determined as to

significant performance levels, a Scheffé post hoc test was applied in order to reveal the clusters that outperformed others or that performed at the same level, even though their performance outcomes have always varied in some way. The results of these 2 tests (i.e., ANOVA and Scheffé post hoc test) are presented in Tables 3 and 4.

4.2.1. Overview of the Service Quality

(1) Criteria with No Significant Differences. The ANOVA results indicate that there were non-significant differences in customer complaints ($F=1.013$, $p=0.421$) in 2015, as shown in Table 3. The last 3 attributes, including customer complaints ($F=0.875$, $p=0.494$), baggage mishandling ($F=2.557$, $p=0.066$), and personal document handling ($F=1.651$, $p=0.196$) in 2016 are shown in Table 4. These figures mean, as in the above-mentioned criteria, all clusters delivered similar service quality levels in both years because the scores they attained were the same, statistically speaking.

(2) Criteria with Significant Differences. Besides non-significant differences, the ANOVA results reveal there were statistically significant differences for the remaining criteria comprised of the check-in process ($F=15.648$, $p<0.001$), the boarding process ($F=29.712$, $p<0.001$), staffing attitude ($F=31.474$, $p<0.001$), baggage mishandling ($F=12.294$, $p<0.001$) (at a 1% alpha level), and personal document mishandling ($F=4.124$, $p<0.012$) (at a 5% alpha level) (see Table 3), followed by the check-in process ($F=35.825$, $p<0.001$), boarding process ($F=36.176$, $p<0.001$), and the staffing attitude ($F=25.485$, $p<0.001$) (see Table 4). They have indicated where the differential ground handling service quality was in both years. In 2015, the service quality of the check-in process, the boarding process, the staffing attitude, baggage mishandling, and personal documents mishandling provided to passengers was considerably different, as accorded by the 5 clusters. In other words, there existed substantial gaps of performance among the 5 clusters. Similarly, in 2016, statistically significant differences are found in the check-in process, the boarding process, and staffing attitude. It is obvious that among these 3 criteria, the performance distances of the 3 clusters were considerably vast.

The 2 years of 2015 and 2016 present dissimilar performance values, as localized in some certain criteria. As such, it is believed that more managerial attention needs to be focused on these criteria. In theory, ANOVA assists to identify the criteria for which there is statistically significant difference of at least one pair of clusters, but does not indicate where the differences may lie [51]. In order to figure out the distinct differences of performance among the clusters in these individual criteria, the authors used the Scheffé post hoc test. In this regard, the benchmarks and the underperforming clusters in the designated criteria could be determined in the next part.

4.2.2. Benchmarks Revealed from Comparison of Clusters' Service Quality. Tables 3 and 4 show the results of Scheffé

TABLE 3: Results of one-way ANOVA and Scheffé post hoc test for 2015.

Criteria	Cluster means in each criterion	ANOVA results	Post hoc results
Check-in process	A1 = 78.7550 A2 = 82.8733 A3 = 81.3200 A4 = 85.8100 A5 = 84.6333	$F = 15.648 > F_{crit} (4.23, 0.05) = 2.80$ Sig = 0.000***	A4, A5 > A3; A2 > A1
Boarding process	A1 = 77.8875 A2 = 78.7900 A3 = 82.8025 A4 = 83.8625 A5 = 83.2767	$F = 29.712 > F_{crit} (4.23, 0.05) = 2.80$ Sig = 0.000***	A4, A5, A3 > A2, A1
Staffing attitude	A1 = 78.8875 A2 = 81.1400 A3 = 81.7963 A4 = 85.3600 A5 = 84.0350	$F = 31.474 > F_{crit} (4.23, 0.05) = 2.80$ Sig = 0.000***	A4, A5 > A2, A3 > A1
Customer complaints	A1 = 0.0925 A2 = 0.1217 A3 = 0.0563 A4 = 0.1700 A5 = 0.0583	$F = 1.013 < F_{crit} (4.23, 0.05) = 2.80$ Sig = 0.421	Non-significant differences
Baggage mishandling	A1 = 2.4850 A2 = 0.2950 A3 = 1.1475 A4 = 5.9050 A5 = 1.1050	$F = 12.294 > F_{crit} (4.23, 0.05) = 2.80$ Sig = 0.000***	A2, A5, A3, A1 > A4
Personal document mishandling	A1 = 0.5525 A2 = 0.6683 A3 = 0.3437 A4 = 1.3775 A5 = 0.2550	$F = 4.124 > F_{crit} (4.23, 0.05) = 2.80$ Sig = 0.012**	A5, A3, A1, A2 > A4

Cluster 1 = A1; cluster 2 = A2; cluster 3 = A3; cluster 4 = A4; cluster 5 = A5. *** $p < 0.001$; ** $p < 0.05$.

post hoc test and present an overview of service quality of the network in 2015 and 2016. By means of this test, real cluster performance levels are determinable that lead to benchmark formulation.

The detailed descriptions for 5 criteria with significant differences are as follows:

- (i) *Check-in process*: A1 is seen to be the worst/least performance cluster according to the criterion. Airports in this cluster kept passengers waiting in lengthy check-in queues, as well as beyond average check-in process times. A2 is better than A1. There were no different gaps of performance between A4 and A5. Both clusters exceeded performance standards and passengers appeared satisfied with the quality of their check-in service and controlled passengers' wait times in queue. However, A5 met more targets stipulated by VNA than did A4. Hence, A4 and A5 should become the benchmark for A3, while A2 can become the benchmark for A1.

In 2016, statistically speaking, with the formation of 4 significant service levels, the criterion had the most performance levels among the 5 clusters. Regarding these levels, B3 offered a similar service quality level with B1. They needed improvement

because they were formed by the least performing airports, especially they included 4 out of total 5 Japanese airports where passengers expected to be served at a higher service quality. Their check-in service quality was below the flyers' expectations. Also, 2 pairs of clusters B3 and B5; B5 and B2 were of equal performance levels. In contrast, B4 was able to handle this exceptional service standard, and it became the benchmark for B5 and B2, while B5 should become the benchmark for B1, and B2 was benchmark for B3.

- (ii) *Boarding process*: the same performance level occurred between A1 and A2 and among A3, A5, and A4. However, A3, A5, and A4 outperformed A1 and A2. Both clusters did not provide the convenience of a smooth boarding process. The information given at the boarding gates remained unclear. The speech for boarding was sometimes unclear which caused confusion among passengers. Likewise, the staff did not promptly render assistance which met passengers' requirements. Thus, one benchmark is suggested.

Unlikely, the findings revealed that there were no differences of service quality rendered between B4 and B5; B3 and B1; and B1 and B2. In fact, significant

TABLE 4: Results of one-way ANOVA and Scheffé post hoc test for 2016.

Criteria	Cluster means in each criterion	ANOVA results	Post hoc results
Check-in process	B1 = 76.7500 B2 = 82.3300 B3 = 79.1700 B4 = 88.0000 B5 = 81.9200	$F = 35.825 > F_{crit} (4.23, 0.05) = 2.80$ Sig = 0.000***	$B4 > B5 > B1; B4 > B2; B2 > B3$
Boarding process	B1 = 76.0000 B2 = 75.0000 B3 = 78.3300 B4 = 83.0000 B5 = 81.4600	$F = 36.176 > F_{crit} (4.23, 0.05) = 2.80$ Sig = 0.000***	$B4, B5 > B3 > B2; B5 > B1$
Staffing attitude	B1 = 77.3475 B2 = 80.8300 B3 = 79.4283 B4 = 86.2000 B5 = 82.2546	$F = 25.485 > F_{crit} (4.23, 0.05) = 2.80$ Sig = 0.000***	$B4 > B5, B2, B3; B4 > B3; B2, B5 > B1$
Customer complaints	B1 = 0.0800 B2 = 0.0767 B3 = 1.2233 B4 = 0.1150 B5 = 0.0031	$F = 0.875 < (4.23, 0.05) = 2.80$ Sig = 0.494	Non-significant differences
Baggage mishandling	B1 = 0.1875 B2 = 1.0467 B3 = 0.8800 B4 = 5.1000 B5 = 1.7262	$F = 2.557 < F_{crit} (4.23, 0.05) = 2.80$ Sig = 0.066	Non-significant differences
Personal document mishandling	B1 = 0.2150 B2 = 0.8200 B3 = 0.4350 B4 = 1.0200 B5 = 0.3254	$F = 1.651 < F_{crit} (4.23, 0.05) = 2.80$ Sig = 0.196	Non-significant differences

Cluster 1 = B1; cluster 2 = B2; cluster 3 = B3; cluster 4 = B4; cluster 5 = B5. *** $p < 0.001$.

performance gaps existed between B2 and B3; between B3 and B5 or B4; and between B1 and B5. B4 and B5 were the front runners in handling boarding tasks, while B3, B1, and B2 needed to demonstrate better control of this process. B4, B5, and B3 turned out to be benchmarks for the following lower performance clusters.

- (iii) *Staffing attitude*: the productivity of A4 and A5 and A2 and A3 did not significantly differ. There appeared two substantial gaps. A1 needs to solve the problem in communication between the staffing and passengers. Staff should deliver information clearly along with an appearance of helpfulness and friendliness. Based on the two gaps, A4 and A5 can be seen as benchmarks for A3 or A2, while A2 can be A1's benchmark.

The 2016 results presented 3 significant levels because of the same performance among B5, B2, and B3 and between B3 and B1. Benchmarks therefore can be seen as B4, B2, or B5.

- (iv) *Baggage mishandling*: in comparison of 5 clusters in 2015, there were no gaps among A2, A5, A3, and A1, but relatively big gaps among them and A4. Therefore, they can become the benchmarks for A4.

The airports in A4 had the lowest scores. There were serious problems evident with misrouted baggage, damaged baggage, wrongly tagged baggage, and wrongly loaded baggage. The number of mishandled baggage cases at this airport was the highest recorded in the system.

- (v) *Personal document mishandling*: The results indicate that the performance of A5 was statistically equal to that of A3, A1, and A2, while A4 represents the worst performing cluster. A4's check-in agents made errors in controlling passenger's travel documents in direct noncompliance with the entry, transit, or other requirements of the countries of travel. Hence, any of A2, A3, A4, and A5 can become the benchmarks for A1.

To sum up, these two tests may provide VNA's administrators with a criteria-based insight coupled with an underperforming service quality directly affecting VNA passengers throughout the international sphere of operations. For each of these criteria, managers had to recognize the real service-level performance of each airport clusters. Airline management will have the potential to identify outperforming clusters for which pro-offered incentives

encourage and maintain qualitative service excellence. Managers can also define given performance according to a variety of parameters for which action is called for to improve them. It is believed that test results will support VNA's strategic management to direct effective and efficient ground handling performance. Furthermore, these findings may also serve station managers in charge of service provision in the airports, so they can draw comparisons as to the productivity level of their airport with other airports that are located in their cluster.

Identifying the significantly different levels of service performance among the clusters is of great managerial importance since the specific performance level of the next higher clusters can be easily used as a benchmark (suggested target) for the one immediately following. In this way, VNA management is now able to establish a more realistic and readily achievable set of targets for certain clusters in the future.

5. Conclusion and Recommendation

The present paper has used a model of cluster analysis, ANOVA, and Scheffé post hoc to better understand the service quality of VNA's ground handling service as it was recently delivered to VNA passengers, at 28 international airports, regarding 6 criteria. The method used to find benchmarks for the next service period for general improvement has been demonstrated. This analysis aims to support the company's overall management strategy to attract repeat business by focusing on customer satisfaction potential leading to profitability.

The study findings have presented an overview of the service performance level of this organization which is beneficial to VNA's top management, as well as to station managers. Top management may realize the service system's weaker areas to afford improvement, benchmarking, and potential targets evident in the next evaluative periods. Through the analysis, 5 classified airport clusters were obtained in 2015. These clusters actually possessed significant differences in 5 of the criteria, namely, the check-in process, the boarding process, staffing attitude, baggage mishandling, and personal document mishandling. The 5 classified clusters obtained in 2016 possessed 3 criteria, namely, the check-in process, the boarding process, and staffing attitude. For these individual criteria, the significant differences in productivity levels among these clusters have been simply identified. Subsequently, the outperforming clusters and the underperforming clusters are most revealed. It is recommended that the outperforming clusters become the benchmark for those clusters which are immediately antecedent. As such, the scores of the benchmark clusters may also suggest realistic and objective targets for the coming service period.

More importantly, by determining the meaningfully different levels of service quality located among the clusters, this outcome provides the expatriate station managers with a clearer viewpoint of the localized productivity level performed in relation to other airports within their own cluster.

Using the cluster analysis method to obtain benchmarks as achievable targets will advance in understanding the quality of the service system so that further recommendations may be meted out to VNA's senior managers and VNA's international airport representatives, as follows.

Firstly, in securing representative ground handling companies, VNA faces a distinct loss of management control [43]. In order to minimize this risk, the company has been implementing SLA through the supervisory intercession of expatriate station managers who can vet desired service quality. Therefore, the selection of station managers, the use of an effective motivational program, and the application of state-of-the-art check-in technologies are of key importance.

Station managers who are selected should be experienced and skillful since they play an important role in building partnership, as well as in coordinating directly with the ground handling agents.

Regarding the motivation of these supervisory representatives, VNA should design an effective managerial program. Such a program should cover the routine meetings of all station managers to insure knowledge sharing, and there should be special attention paid to underperforming airports, made possible through reports on explaining problem-solution activities and justification of events.

Although VNA's online check-in service has been available, VNA should consider more alternatives to using technology-based self-service such as self-check-in kiosks and barcode-activated check-in. These auto-check-in facilities offer numerous advantages to fliers such as enabling them to pick desired seating, reducing check-in times, and preventing congestion and delay at check-in counters [52]. However, passengers' intentions and satisfaction in using them is still negligible. By adopting these technologies, airlines such as VNA may offer additional benefits or incentives for prior seat selection in order to encourage departing passengers to use self-service technologies as well as allocate local staff to provide specialized instruction to assist passengers in the use of these facilities [53]. Moreover, concerning limited space, the choice of the self-check-in kiosk locations proximate to the luggage conveyor system will further enable satisfactory service for both the airline and the fliers [54]. In sum, reducing waiting time is insuring passengers' satisfaction since excessive wait times are considered to be one factor directly affecting overall fliers' satisfaction [53], and check-in wait times have the heaviest weight in the overall passenger service process.

Secondly, the interaction of passengers, distinct airport features, and airport entry-exit related procedures are all crucial to the customer service experience. The station managers should have an awareness of the passenger profiles of those they serve (i.e., nationality, specialized groups, and various traveling purposes) at the airports since these features impact passengers' expectations of relative service quality [24]. Similarly, although it is clear that SLA signed is rather detailed, the different characteristics that individual airports have should be of concern. There are some advantageous and disadvantageous elements, such as employee skill sets, passenger traffic and flow, and aircraft movement and placement. Employees with good knowledge and skills

contribute to a better service quality level [43] and airports with capacity constraints may impose delays on aircraft and passengers [19]. Apart from regular procedures, such as security filters, passport control, and pinch-points that may differ at various airports, there may also be occasional or frequent security features specifically imposed due to threat levels on the passengers at certain airports. The station managers need to ascertain the situation in a timely manner to carry out the necessary actions that may assist passengers.

5.1. Suggested Research. Future research should focus on a comparison of the ground handling service quality management methods of VNA next to other carriers and how to generate service quality targets in groups which would support the positive assessment of VNA ground handling service worldwide.

Data Availability

Access to data is restricted due to commercial confidentiality.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] B. S. Everitt, S. Landau, M. Leese, and D. Stahl, *Cluster Analysis*, King's College, London, UK, 5th edition, 2011.
- [2] J. F. Hair Jr, M. Wolfinger, A. H. Money, P. Samouel, and M. J. Page, *Essentials of Business Research Methods*, Routledge, Abingdon, UK, 2016.
- [3] G. Punj and D. W. Stewart, "Cluster analysis in marketing research: review and suggestions for application," *Journal of Marketing Research*, vol. 20, no. 2, pp. 134–148, 1983.
- [4] J. H. Ward, "Hierarchical grouping to optimize an objective function," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 236–244, 1963.
- [5] L. Davison and T. Ryley, "Tourism destination preferences of low-cost airline users in the East Midlands," *Journal of Transport Geography*, vol. 18, no. 3, pp. 458–465, 2010.
- [6] P. Malighetti, S. Pailari, and R. Redondi, "Airport classification and functionality within the European network," *Problem Perspective Management*, vol. 7, no. 1, pp. 183–196, 2009.
- [7] R. Mayer, "Airport classification based on cargo characteristics," *Journal of Transport Geography*, vol. 54, pp. 53–65, 2016.
- [8] H.-A. Vogel and A. Graham, "Devising airport groupings for financial benchmarking," *Journal of Air Transport Management*, vol. 30, pp. 32–38, 2013.
- [9] S. Hands and B. Everitt, "A Monte Carlo study of the recovery of cluster structure in binary data by hierarchical clustering techniques," *Multivariate Behavioral Research*, vol. 22, no. 2, pp. 235–243, 1987.
- [10] V. Adikariwattage, A. G. de Barros, S. C. Wirasinghe, and J. Ruwanpura, "Airport classification criteria based on passenger characteristics and terminal size," *Journal of Air Transport Management*, vol. 24, pp. 36–41, 2012.
- [11] G. Burghouwt and R. Redondi, "Connectivity in air transport networks. An assessment of models and applications," *Journal of Transport Economics and Policy*, vol. 47, pp. 35–53, 2013.
- [12] H. Huber, "Spatial structure and network behaviour of strategic airline groups: a comparison between Europe and the United States," *Transport Policy*, vol. 16, no. 4, pp. 151–162, 2009.
- [13] H. Rodríguez-Déniz, P. Suau-Sanchez, and A. Voltes-Dorta, "Classifying airports according to their hub dimensions: an application to the US domestic network," *Journal of Transport Geography*, vol. 33, pp. 188–195, 2013.
- [14] J. Sarkis and S. Talluri, "Performance based clustering for benchmarking of US airports," *Transportation Research Part A: Policy and Practice*, vol. 38, no. 5, pp. 329–346, 2004.
- [15] A. Jessop, "A decision aid for finding performance groups," *Benchmarking: An International Journal*, vol. 19, no. 3, pp. 325–339, 2012.
- [16] H. Rodríguez-Déniz and A. Voltes-Dorta, "A frontier-based hierarchical clustering for airport efficiency benchmarking," *Benchmarking: An International Journal*, vol. 21, no. 4, pp. 486–508, 2014.
- [17] Y.-C. Chiou and Y.-H. Chen, "Route-based performance evaluation of Taiwanese domestic airlines using data envelopment analysis," *Transportation Research Part E: Logistics and Transportation Review*, vol. 42, no. 2, pp. 116–127, 2006.
- [18] C. W. Hill, G. R. Jones, and M. A. Schilling, *Strategic Management, Theory*, Cengage Learning, Boston, MA, USA, 12th edition, 2016.
- [19] R. Etemad-Sajadi, S. A. Way, and L. Bohrer, "Airline passenger loyalty: the distinct effects of airline passenger perceived pre-flight and in-flight service quality," *Cornell Hospitality Quarterly*, vol. 57, no. 2, pp. 219–225, 2016.
- [20] B. Gale, *Monitoring Customer Satisfaction and Market-Perceived Quality (American Marketing Association Worth Repeating Series No. 922CSO1)*, American Marketing Association, Chicago, IL, USA, 1992.
- [21] R. Buzzell and B. Gale, *The PIMS Principle: Linking Strategy to Performance*, Free Press, New York, NY, USA, 1987.
- [22] M. D. Johnson, G. Nader, and C. Fornell, "Expectations, perceived performance, and customer satisfaction for a complex service: the case of bank loans," *Journal of Economic Psychology*, vol. 17, no. 2, pp. 163–182, 1996.
- [23] F.-Y. Chen and Y.-H. Chang, "Examining airline service quality from a process perspective," *Journal of Air Transport Management*, vol. 11, no. 2, pp. 79–87, 2005.
- [24] D. Gilbert and R. K. C. Wong, "Passenger expectations and airline services: a Hong Kong based study," *Tourism Management*, vol. 24, no. 5, pp. 519–532, 2003.
- [25] A. Parasuraman, V. A. Zeithaml, and L. L. Berry, "SERVQUAL: a multiple-item scale for measuring consumer perception of service quality," *Journal of Retailing*, vol. 64, no. 1, p. 12, 1988.
- [26] F. Sultan and M. C. Simpson, "International service variants: airline passenger expectations and perceptions of service quality," *Journal of Services Marketing*, vol. 14, no. 3, pp. 188–216, 2000.
- [27] R. Wang, Y. H. Shu-Li, M. L. Hsu, Y. H. Lin, and M.-L. Tseng, "Evaluation of customer perceptions on airline service quality in uncertainty," *Procedia—Social and Behavioral Sciences*, vol. 25, pp. 419–437, 2011.
- [28] H.-C. Wu and C.-C. Cheng, "A hierarchical model of service quality in the airline industry," *Journal of Hospitality and Tourism Management*, vol. 20, pp. 13–22, 2013.

- [29] R.-C. Jou, S.-H. Lam, C.-W. Kuo, and C.-C. Chen, "The asymmetric effects of service quality on passengers' choice of carriers for international air travel," *Journal of Advanced Transportation*, vol. 42, no. 2, pp. 179–208, 2008.
- [30] E. Babakus and G. W. Boller, "An empirical assessment of the SERVQUAL scale," *Journal of Business Research*, vol. 24, no. 3, pp. 253–268, 1992.
- [31] J. J. Cronin and S. A. Taylor, "Measuring service quality: a reexamination and extension," *Journal of Marketing*, vol. 56, no. 3, pp. 55–68, 1992.
- [32] L. F. Cunningham, C. E. Young, and M. Lee, "Perceptions of airline service quality: pre and post 9/11," *Public Works Management & Policy*, vol. 9, no. 1, pp. 10–25, 2004.
- [33] A. Parasuraman, L. L. Berry, and V. A. Zeithaml, "Refinement and reassessment of the SERVQUAL scale," *Journal of Retailing*, vol. 67, no. 4, pp. 420–450, 1991.
- [34] A. Parasuraman, V. A. Zeithaml, and L. L. Berry, "A conceptual model of service quality and its implications for future research," *Journal of Marketing*, vol. 49, no. 4, pp. 41–50, 1985.
- [35] A. Parasuraman, V. A. Zeithaml, and L. L. Berry, "Reassessment of expectations as a comparison standard in measuring service quality: implications for further research," *Journal of Marketing*, vol. 58, no. 1, pp. 111–124, 1994.
- [36] R. K. Teas, "Expectations, performance evaluation, and consumers' perceptions of quality," *Journal of Marketing*, vol. 57, no. 4, pp. 18–34, 1993.
- [37] C.-C. Chou, L.-J. Liu, S.-F. Huang, J.-M. Yih, and T.-C. Han, "An evaluation of airline service quality using the fuzzy weighted SERVQUAL method," *Applied Soft Computing*, vol. 11, no. 2, pp. 2117–2128, 2011.
- [38] H. Gupta, "Evaluating service quality of airline industry using hybrid best worst method and VIKOR," *Journal of Air Transport Management*, vol. 68, pp. 35–47, 2018.
- [39] S.-H. Tsaur, T.-Y. Chang, and C.-H. Yen, "The evaluation of airline service quality by fuzzy MCDM," *Tourism Management*, vol. 23, no. 2, pp. 107–115, 2002.
- [40] Y.-H. Chang and C.-H. Yeh, "A survey analysis of service quality for domestic airlines," *European Journal of Operational Research*, vol. 139, no. 1, pp. 166–177, 2002.
- [41] J. J. H. Liou, C.-C. Hsu, W.-C. Yeh, and R.-H. Lin, "Using a modified grey relation method for improving airline service quality," *Tourism Management*, vol. 32, no. 6, pp. 1381–1388, 2011.
- [42] B. D. Bowen and D. E. Headley, *Airline Quality Rating 2015: The 25th Year Reporting Airline Performance*, Wichita State University, Wichita, KS, USA, 2015.
- [43] C.-C. Hsu and J. J. H. Liou, "An outsourcing provider decision model for the airline industry," *Journal of Air Transport Management*, vol. 28, pp. 40–46, 2013.
- [44] B. Kucukaltan and Y. I. Topcu, "Assessment of key airline selection indicators in a strategic decision model: passengers' perspective," *Journal of Enterprise Information Management*, vol. 32, no. 4, pp. 646–667, 2019.
- [45] M. Mellat-Parast, D. Golmohammadi, K. L. McFadden, and J. W. Miller, "Linking business strategy to service failures and financial performance: empirical evidence from the U.S. domestic airline industry," *Journal of Operations Management*, vol. 38, no. 1, pp. 14–24, 2015.
- [46] L. Zhang, L. Zhang, P. Zhou, and D. Zhou, "A non-additive multiple criteria analysis method for evaluation of airline service quality," *Journal of Air Transport Management*, vol. 47, pp. 154–161, 2015.
- [47] D. Gursoy, M.-H. Chen, and H. J. Kim, "The US airlines relative positioning based on attributes of service quality," *Tourism Management*, vol. 26, no. 1, pp. 57–67, 2005.
- [48] W. Wu, C. L. Wu, T. Feng, H. Zhang, and S. Qiu, "Comparative analysis on propagation effects of flight delays: a case study of China airlines," *Journal of Advanced Transportation*, vol. 2018, Article ID 5236798, 10 pages, 2018.
- [49] A. Toruń, C. Burniak, J. Biały et al., "Challenges for air transport providers in Czech Republic and Poland," *Journal of Advanced Transportation*, vol. 2018, Article ID 6374592, 7 pages, 2018.
- [50] J. Rezaei, O. Kothadiya, L. Tavasszy, and M. Kroesen, "Quality assessment of airline baggage handling systems using SERVQUAL and BWM," *Tourism Management*, vol. 66, pp. 85–93, 2018.
- [51] D. R. Cooper and P. S. Schindler, *Business Research Methods*, McGraw-Hill Irwin, New York, NY, USA, 12th edition, 2014.
- [52] D. Weiss, "Analysis: kiosk uptime revenue," *Airport Business*, vol. 7, pp. 27–29, 2006.
- [53] C.-I. Hsu, C.-C. Chao, and K.-Y. Shih, "Dynamic allocation of check-in facilities and dynamic assignment of passengers at air terminals," *Computers & Industrial Engineering*, vol. 63, no. 2, pp. 410–417, 2012.
- [54] H.-L. Chang and C.-H. Yang, "Do airline self-service check-in kiosks meet the needs of passengers?" *Tourism Management*, vol. 29, no. 5, pp. 980–993, 2008.
- [55] L. Di Pietro, R. Guglielmetti Mugion, F. Musella, M. F. Renzi, and P. Vicard, "Monitoring an airport check-in process by using Bayesian networks," *Transportation Research Part A: Policy and Practice*, vol. 106, pp. 235–247, 2017.