Journal of Air Transport Management 106 (2023) 102305

Contents lists available at ScienceDirect

Journal of Air Transport Management

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

Methodological framework for a deeper understanding of airline profit cycles in the context of disruptive exogenous impacts

Manuel Renold [a], Janik Vollenweider [a], Nemanja Mijović [b], Jovana Kuljanin [c,*], Milica Kalić [b]

[a] Zurich University of Applied Sciences, Technikumstrasse 81, 8400, Winterthur, Switzerland
[b] University of Belgrade – Faculty of Transport and Traffic Engineering, Vojvode Stepe 305, 11 000, Belgrade, Serbia
[c] Polytechnic University of Catalonia, Esteve Terradas 5, 08860, Castelldefels, Barcelona, Spain

A R T I C L E I N F O

Keywords:
Airline profitability cycles
Operating profit prediction
k-means clustering
System dynamic modelling
Principal component analysis

A B S T R A C T

This paper combines the k-means clustering method in combination with PCA and the system dynamic modeling approach to derive a better insight into the behavior of airline profitability during the time span of 1995 until 2020. The model includes various explanatory variables that capture different aspects of airline economic and operational metrics, whose fluctuations may affect the airline profitability. By forecasting these exogenous variables, the system dynamic model is used to predict airline profitability through 2025 and answer the question of whether the US airline industry will return to its pre-COVID 19 pandemic state. The latter research question can be agreed with, as the effect of introducing a fourth dimension derived from Principal Component Analysis (PCA) to sufficiently cover the variation within the dataset during the years of COVID-19 pandemic diminishes towards the end of the forecast period. Furthermore, the key measures from PCA imply that under the assumption of continuous growth and a non-exogenous shock, future years will not cluster in past years. The six different clusters from 2019 to 2025 showed how the system stays in a certain state for a few years and then drifts further to a new state. There are only a few variables that change to transfer from one cluster to the next.

1. Introduction

The airline industry is traditionally characterized as a strongly cyclical business with extremely high capital turnovers, contrasted by lean profit margins (Doganis, 2005). Periods of profitable years have been alternating with severe periods of financial losses, a trend that has been observed over the last seventy years on a global level (Franke, 2007). Furthermore, airlines' return on invested capital (ROIC) is the lowest among other supply chain sectors in air transportation (IATA Economics, 2013). This highlights the inability of the airline industry to sustain financial health, despite a period of increasing passenger numbers, fleet size and overall network growth over recent decades (Maung et al., 2022). The magnitude of growth can be observed in the number of air passengers, which grew from approximately 2.7 billion to 4.5 billion during the decade of 2010–2019 (Statista, 2020).

Although profitability performances did not play a significant role at the dawn of commercial aviation, when state-owned carriers used to be supported by government subsidies and seats (The Economist, 2014), the wave of liberalization has permanently changed the market conditions in several important aspects. First, it fostered the process of privatization which entailed more careful management of the airlines' balance sheet, considering both cost and revenue side. Second, liberalization prompted fierce competition among airlines in terms of fare and level of service offered to passengers. Finally, it led to the emergence of new business models, bringing the low-cost model to the forefront on short- and medium-haul routes in Europe and the U.S. This new environment, entwined with both a structural sensitivity to external shocks (Franke and John, 2011) and unevenly-balanced shares of the global value in the air transport vertical channel (Button and McDougall, 2006; Martini, 2022), had a tremendous impact on poor airline financial performance. On the other hand, cost saving derived from the introduction of new technologies appeared to be beneficial for passenger and market share growth rather than improvements in profits (Maung et al., 2022).

Given its complex nature, airline profitability has been addressed by several past studies from different perspectives, each of them aiming to shed some light on a specific profitability issue. These include the relationship between profitability performance and airlines' choice of business model (e.g., Collins et al., 2011; Alamdari and Fagan, 2005;

* Corresponding author.
E-mail addresses: reno@zhaw.ch (M. Renold), volj@zhaw.ch (J. Vollenweider), nemanja.mijovic91@gmail.com (N. Mijović), jovana.kuljanin@upc.edu (J. Kuljanin), m.kalic@sf.bg.ac.rs (M. Kalić).

https://doi.org/10.1016/j.jairtraman.2022.102305
Received 8 December 2021; Received in revised form 10 August 2022; Accepted 20 September 2022
Available online 1 October 2022
0969-6997/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Franke and John, 2011); the implications of different operational decision and strategies (e.g., Chin and Tay, 2001; Barnhart et al., 2009), the relationship between profit fluctuations and industry value chain (IATA Economics, 2013) and the cyclical behavior of airline profits (Jiang and Hansman, 2004; Pierson and Sterman, 2013). Unlike numerous studies performed in each of these respective areas, the behavior of airline profit cycles has been tackled by only a small number of scholars. For instance, Jiang and Hansman (2004) used the spectrum analysis and determined that aggregate earnings of the U.S airline industry have fluctuated with an average peak-to-peak period of approximately 10 years. In addition to spectrum analysis, this type of study mainly employs system dynamic modeling, which is proven to provide more reliable forecasts of short-to mid-term trends than statistical models (Lyneis, 2000). For instance, Liehr et al. (2001) combined the system dynamic approach with a statistical forecasting model to analyze cycle-generating structures in the airline market. More specifically, the study aimed to support Lufthansa Airlines to identify alternative strategies for effective management of cycle behavior. Expanding the boundary of the models proposed by Lyneis (2000) and Liehr et al. (2001), Pierson and Sterman (2013) included an endogenous account of feedback omitted from earlier work, considering price setting, wages, and air travel demand. The authors found substantial evidence that aggressive use of yield management—varying prices to ensure high load factors (capacity utilization)—may have the unintended effect of increasing earnings variance. One of the recent studies performed in this field was conducted by Cronrath (2018), showing satisfactory results in terms of model accuracy, considering the complex interrelationship between the economic and operational data up to 2010. However, after this period, the predictive capability of the model significantly decreases (i.e. estimated values deviate from the actual ones).

Furthermore, with COVID-19 being one of the most detrimental exogenous impacts faced by the airline industry, the simulation performed by using the Cronrath (2018) model in its original form has shown that it is not able to represent this crisis. The COVID-19 crisis was simulated in the model as an exogenous shock, reducing revenue passenger miles (RPM) by 48 percent in 2020 and 2021. Without changes to the model, the entire system collapses. This can be seen in the fact that RPM and available seat miles (ASM) drop to zero and do not recover in the simulated period until 2025. It is presumed that the high accuracy until 2010, the decreased accuracy until 2020 and the failure in simulating the COVID19-crisis come from overfitting the training data. The purpose of this paper is to extend the previous work of Cronrath (2018) by proposing a novel approach that combines system dynamic model and k-means cluster analysis taking into account the specific effect of COVID-19 on airline net profit cycles. To the best of our knowledge, the combination of system dynamic model and k-means cluster analysis has not been used in the previous studies that investigate the airline profit cycles. Having in mind that the adaptation of the Cronrath model imposed significant challenges due to the large number of variables and their complex interdependencies, a new and simpler system dynamic model was created, focusing only on a few core processes and using more exogenous variables. The exogenous variables are primarily forecast estimates obtained from relevant aviation and financial organizations (IATA reports, Deloitte, etc.), as well as the actual data for the year 2020. By combining the system dynamics and k-means clustering algorithm, the paper aimed at the integration of two different methods for analyzing and dealing with the issue of airline industry profit behavior. The proposed methodology provides the possibility to examine the problem from two complementing perspectives. With the given approach, the paper aims to address the question whether an exogenous impact such as COVID-19 would lead to the paradigm shift in airline profitability cycle behavior.

Although the available reports already demonstrate that traffic numbers are steadily returning to pre-COVID-19 levels (EUROCONTROL, 2022), the question of profitable growth still remains vague. Airlines worldwide have been caught flat-footed by the crisis, having in mind the 10-year positive trend in airline net profits, stable geopolitical situations, and persistently rising demand for air travel. For instance, by the end of 2019, the U.S. airline industry experienced historic levels of profitability, with the six largest U.S.-based airlines (Delta, American, United, Southwest, Alaskan and JetBlue) having combined revenues of more than $175 billion and combined operating incomes of almost $19 billion with very manageable debt levels (Shaked and OrelowItz, 2020). Nevertheless, the impact of the pandemic was so severe that airlines' positive financial performances started to dramatically melt away, and shortly after its onset, loss of gross operating revenue was estimated to be between 112 and 135 billion USD (ICAO, 2020). Though airline operating profit forecasts have generally been viewed as a complex task to perform (Cronrath, 2018), the speed and breadth of the impact of COVID-19 has imposed a new level of uncertainty and challenges for modeling analysts.

The aim of this study is to shed some light on the potential recovery pattern and the behavior of the net profits in the upcoming years. This will have particular implications to aviation authorities and airline managements to better understand the effect of the exogenous shocks on airline profit cycles and anticipate important structural changes in the industry. By analyzing various explanatory variables and its contribution to past and future years, the main driver related to airline operating profit behavior will be identified. Knowledge of the contribution and interaction of the explanatory variables will enable the aforementioned aviation stakeholder to compare future years with past system states.

The paper is organized as follows: Section 2 provides a review of literature on profitability in the airline industry by emphasizing different methods and techniques that were used to tackle this complex issue. Section 3 describes the methodological framework proposed in the paper and a brief explanation of the k-means cluster method and the system dynamics modeling approach and how they are combined in the paper. Further, the data used as exogenous and explanatory variables, including sources for past years and forecasts, are documented. In Section 4, a two-stage cluster analysis is performed using the main explanatory variables from the predicted variables and the output of the system dynamics model. Furthermore, the resulting clusters are analyzed in more detail by focusing on the underlying key measures of principal component analysis (PCA). Section 5 concludes the paper with a summary of major findings and an outlook of potential future research.

2. Literature review

The airline profit cycle is very often associated with the underlying economic cycle as they typically fluctuate in the same direction (Lenoir, 1998). The period of boom followed by turmoil and recession of the economy has become a common phenomenon across the world and many airlines monitor the development of the economic trends in their main markets. Among a plethora of economic factors that influence the travel demand, the effect of GDP has been extensively investigated in a great number of studies. Applying multiple regression analysis, Chin and Tay (2001) analyzed the profitability of Asian airlines and found that airline profits are positively related to both total gross domestic product and load factor. The authors stressed that higher load factor indicates there is no over-capacity and fuller planes means higher profits. Tarry (2015) stressed that economic growth (measured through GDP growth) does not substantially impact fares, but it should be included in the traffic forecasting models since there is an underlying relation between GDP and air travel demand. The study of Profillidis and Botzoris (2015) acknowledged the importance of economic activity (i.e., GDP per capita) in air travel demand, although they claim that the magnitude of correlation depends on the maturity of markets, with correlation being more pronounced in more mature markets.

However, the effect of economic growth on airline profitability needs to be considered together with airline operating strategy on (capacity) investment. Namely, the airline decision on acquiring a new aircraft typically coincides with the period of favorable economic conditions and high demand for air travel. Given the fact that manufacturers produce the aircraft according to the orders received, it takes typically two to three years for the aircraft to be delivered depending on the aircraft size (Pierson and Sterman, 2013). Past experience has shown that airlines receive the planes, which tend to increase the available capacity at the markets or even to lead to over-capacity, at the time when the demand for travel declines as a result of economic downturn. Consequently, in order to retain their market shares, the airlines lower the fares, which in turn results in diminishing the yield and profits. Thus, the capacity extension presents one the most determining factors for an airline's successful financial performance. Until now, only a small number of studies addressed the relationship between airline profitability and airline growth. The study of Lau and Mattheiss (1992) applied a discrete-time stochastic process (i.e., a finite Markov chain) to analyze and forecast the dynamic structure of the U.S. airline industry in terms of growth in asset size (AST) and net profit margins (NPM). The results eventually showed that the airlines' AST is a key determinant for the survival prospect of airlines - larger airlines (such as Delta, United, American, etc) are, on average, more profitable than smaller airlines (such as Tower, Hawaiian, Frontier and Miami Air). Not surprisingly, the study found that profitable airlines tend to stay profitable for the next transition year. The similar methodology was applied by Chin and Tay (2001) who examined the relationship between assets growth and profitability for Asian airlines. Similarly to the findings in Lau and Mattheiss (1992), the study indicated that the survival probabilities of airlines increase as asset size and profit increase. Both studies emphasized the importance of prompt responsiveness to the changing environment along with the capacity flexibility as decisive factors in financial sustainability.

A northworthy point about profitability cycles is not so much the particular magnitude of one year's performance, but rather the longevity of the cycle and the average growth levels achieved (Hätty and Hollmeier, 2003). The earlier studies showed that the cycle length is endogenously driven indicating several main causes for such behavior - slumping in demand, airlines' inability to adjust their capacity to market conditions, inadequate pricing strategy and excessive costs mainly driven by increasing price of jet oil and skilled labor. In addition to these endogenously driven factors, over the last four decades, the airline industry faced a great number of financial shocks, starting with those arising from the Middle East unrest in the late 70s and early 80s. Moreover, the bursting of the "dot.com" bubble at the onset of the new millennium, combined with the terrorist attack on the Twin Tower in New York on September 11th, 2001 resulted in a "perfect storm" (Franke and John, 2011). The global demand for air travel collapsed immediately causing a substantial financial loss to legacy carriers. Contrary to their full-service competitors, the low cost carriers reaped the benefit of the new business model which enabled them to significantly reduce the cost on one side, and increase the yields on the other. Moreover, the SARS epidemic in East Asia which began in February 2003 and the global economic crisis that occurred at the end of 2008 greatly affected the shape of the profit cycle. Until 2008, it was a prevailing notion that only endogenous causes directly affect the fundamental cycle period, while the exogenous factors were responsible for the change in oscillation amplitude (e.g., Jiang and Hansman, 2004). The post-2008 economic crisis had a particularly severe impact that permanently distorted the fundamental period of profit cycle for the world airline industry that previously accounted for 11 years (Mijović et al., 2018). However, not all world regions had been equally hit by the 2008 economic crisis. Dobruszkes and Van Hamme (2011) pointed out that the crisis had much more affected the USA, Europe and Japan than the rest of the world. Employing the regression analysis, the authors found that the change in the supply of seats is highly dependent on economic growth, confirming the cyclical nature of the airline sector. Being already well-equipped with the lessons learned from the events of 2001/2003, the airlines reacted very quickly as early as 2008 by grounding considerable capacity for the short term. However, the industry followed the so-called "U-shaped" recovery pattern with 2010 being the first year with positive profits after the economic crisis (see Fig. 3). Since then, the changing environment accompanied with emerging business models, improvement in technology and further relaxation of regulation led to the unexpectedly high profits of ten consecutive years.

Among all external shocks, the recent COVID-19 pandemic is foreseen to be one of the most transformative events in the recent past, appearing to radically change the course of the airline business by bringing the entire industry into a frequently cited "new normal'' competitive landscape, whatever this might look like. As of 24 March 2020, many airlines have been brought to a complete stop and, to make matters worse, the recovery pattern for COVID-19 is turning out to be highly uncertain and substantially different than the short-sharp V-shaped pattern observed after the SARS outbreak (Suau-Sanchez et al., 2020). Airlines worldwide have been caught flat-footed by the crisis, having in mind the 10-year positive trend in airline net profits, stable geopolitical situations, and persistently rising demand for air travel. For instance, by the end of 2019, the U.S. airline industry experienced historic levels of profitability, with the six largest U.S.-based airlines (Delta, American, United, Southwest, Alaskan and JetBlue) having combined revenues of more than $175 billion and combined operating incomes of almost $19 billion with very manageable debt levels (Shaked and OrelowItz, 2020). Nevertheless, the impact of the pandemic was so severe that airlines' positive financial performances started to dramatically melt away, and shortly after its onset, loss of gross operating revenue was estimated to be between 112 and 135 billion USD (ICAO, 2020). Though airline operating profit forecasts have generally been viewed as a complex task to perform (Cronrath, 2018), the speed and breadth of the impact of COVID-19 has imposed a new level of uncertainty and challenges for modeling analysts.

3. Methodology and approach

3.1. Conceptual framework

The methodology developed in this paper aims at combining two approaches to derive a better understanding of the future behavior of airline profitability performance (see Fig. 1). The first approach aims at understanding the complex behavior of the airline industry under the exogenous forces e.g., COVID-19 by performing an analysis of the already existing operational and economic data for the U.S. market. The timespan from 1995 until 2020 and 1995 until 2025 were compared using the k-mean clustering method. By clustering the airline financial performance across the years, this paper contributes to an understanding of the dynamics of operating profits of the airline industry. To capture the potential changes in operating profit patterns and driving variables during the airline profit cycles, the study applies the k-means clustering technique and the underlying PCA on the selected explanatory variables, to easily represent high dimensional data on two-dimensional plots and showing the contributions of financial and operational performance during the different cluster before and after the COVID-19 pandemic. The goal of the PCA method in combination with k-means clustering is to identify which variables exert the greatest explanatory power at a certain period of time (years in this case). The paper identifies a set of explanatory variables as important determinants of airline profitability building upon the extensive airline economics and financial performance literature, as well as relevant academic literature (e.g., Maung et al., 2022; Choi et al., 2019).

The second approach, a system dynamics model, comes into play for predicting the airlines operating profit, which is consequently used as an input to k-means to cluster the data set ranging from 1995 to 2020. The airline operating profit is a key variable influenced by many exogenous and endogenous variables such as fuel, labor force, maintenance, passenger and cargo business, and other ancillary businesses (Miranda, 2015). There are already several models in the field of system dynamics, control theory and others that share the same goal of predicting airline

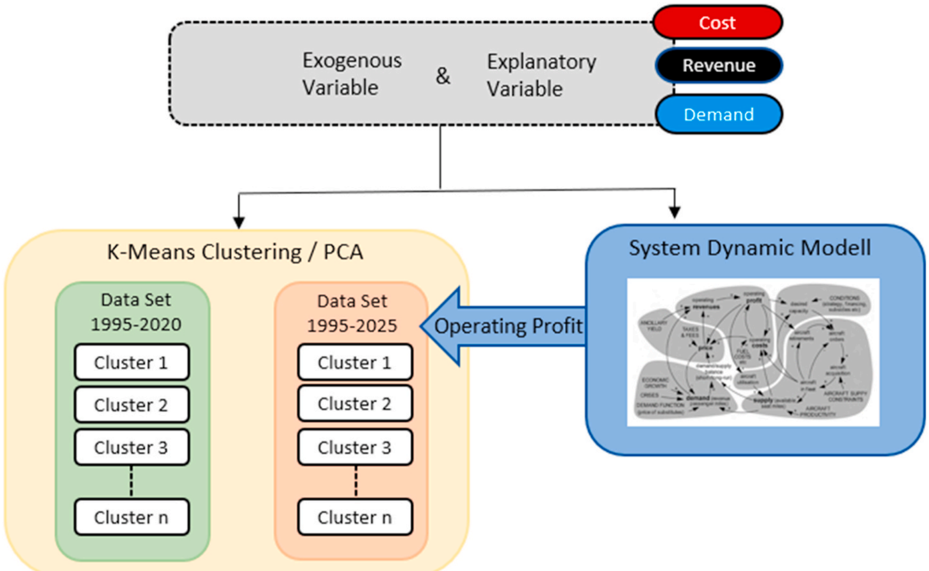


Fig. 1. The overview of the concept.

profits. An extensive literature review on the behavior of the airline profit cycle and its underlying variables was already summarized by Cronrath (2018) in the book titled "The Airline Profit Cycle". In addition, the author proposed a new system dynamic model with the most satisfactory results in terms of model accuracy, considering the complex interrelationship between the economic and operational data up to the year 2010. After this period, the predictive capability of the model significantly decreases (i.e. estimated values deviate from the actual ones). Furthermore, the simulation of the COVID-19 crisis has shown that Cronrath's model in its original form is no longer capable of accurately mapping the profit cycle in the airline industry. It is presumed that the high model accuracy until 2010, the decreased accuracy until 2020 and the failure in simulating the COVID19-crisis comes from overfitting the training data. The required adaptation of the model for further usage revealed its biggest disadvantage, the high complexity of the model due to its strong interdependencies. Thus, a new and simpler system dynamic model is created, focusing only on a few core processes by reducing the number of endogenous variables and increasing the number of exogenous ones.

3.2. Data

For the scope of the k-means clustering algorithm and the system dynamic airline profit model, the study gathered longitudinal aviation operational data from the Airline Data Project (ADP) established by the MIT Global Airline Industry Program (MIT, 2020).

The data used for this analysis is based on the performance of the U.S. airline industry and is sourced from the U.S. Department of Transportation's Form 41 data product. The ADP collection contains a large set of metrics split into seven different categories, reflecting the operational and financial aspects of the airline industry performance, and providing different levels of data aggregation. To enrich our forecasting specification, the study used the IATA Economic Performance of the Airline Industry (IATA, 2020), and the IATA Outlook for Air Transport and the Airline Industry (AGM, 2020) as the main source for the future estimates of Revenue Passenger Miles (RPM) and Load Factor (LF). Table 1 lists the data used here and its source, divided into the information on the data source for the time series (year 1995–2020) and the forecast (year 2021–2025). The data in the corresponding table is used directly as an exogenous variable for the system dynamics model and as an explanatory variable for the k-means clustering approach, or it is used to derive additional operational ratios/metrics. In the absence of literature containing information on forecasts of yield per passenger, average wages or ASM produced per employee, informal interviews were conducted with aviation experts. The experts present the individuals who are mainly engaged in the operational sector of the airline industry with extensive knowledge and experience as well as with the individuals from academia. The interviews were based on open discussion that allowed for the respondents to elaborate in an open manner on the issues of profitability and its recovery pattern in the context of COVID-19. Each interview lasted between one to two hours and was conducted during the first quartal of 2021. Due to requests for confidentiality, the names and roles of the respondents were kept anonymous.

Table 1
Data used for the exogenous variables, where SD stands for system dynamic model, KM for K-means clustering and its underlying PCA and AM for variables used to derive additional metrics and ratios.

Variable	Source: Time Series	Source Forecast	Usage
Operating Profit	MIT (2020)	System Dynamic Model	SD, KM
Revenue Passenger Miles (RPM)	MIT (2020)	IATA (2021b)	SD, KM, AM
Load Factor (LF)	MIT (2020)	IATA (2020a)	SD, KM
ASM per Employee	MIT (2020)	Expert Judgment	SD
Yield per PAX	MIT (2020)	Expert Judgment	SD, KM
Ancillary Yield	MIT (2020)		AM
Total Gallons of Fuel	MIT (2020)		AM
Fuel Price p. Gallon	MIT (2020)	Deloitte (2020b)	SD, KM
Fuel Expenses	MIT (2020)		AM
Average Wage	MIT (2020)	Expert Judgment	SD, KM, AM
Other Expenses	MIT (2020)		KM, AM

The fuel price trend in the past is mainly influenced by the crude oil price trend. A regression analysis showed a highly significant relationship between crude oil price and fuel price with an R^2 of 0.98. Therefore, the crude oil price as an explanatory variable is used to predict the fuel price, based on the data taken from (Deloitte, 2020b). The development of fuel price per gallon from 1995 until 2020 as well as the forecast up to 2025 can be seen in Fig. 3. The variables which do not have specified source of forecast in Table 1 were used to create derived ratios, which were then used as exogenous variables in the models. The basic idea in creating the ratios was to have derived variables whose past behavior is close to a linear or constant behavior and thus, easier to predict up to the year 2025. For more information on these derived ratios, see Section 3.3 *System Dynamic Modeling*. The entire set of explanatory variables and their values can be found in the Appendix.

Table 4 in the Appendix contains all input variables used for the airline profit model and the k-mean clustering approach. Data from 1995 to 2020 are based on existing data, thereafter, are based on reports and expertise,.

3.3. K-means clustering method

The first part of the methodology is based on the application of the k-means clustering algorithm. The method presents one of the most popular unsupervised learning techniques due to its simplicity and efficiency. The algorithm groups similar years in terms of performance profiles, here called "pseudo Steady State of the System" (pSSoS), into clusters based on the selection of the explanatory variables. The pSSoS are created through the application of PCA, where new axes (principal components) are fitted to the data, allowing the data to be presented in such a way that certain dimensions can be removed with minimal loss of information in the data. It will further facilitate the interpretation of each cluster obtained together with the main driving forces (the explanatory variables). The PCA has already found its broad application in different areas of airline industry - airline network structure (Roucolle et al., 2020), environmental efficiency performance (Elhmoud et al., 2021), airport assessment from airlines perspective (Adler and Berechman, 2001) and many others.

The explanatory variables used for k-means clustering are a subset of the variables from Table 1. Specifically, these are operating profit, yield per PAX, RPM, LF, average wages, fuel price per gallon and other expenses.

The cluster analysis consists of several stages: the data collection, a preliminary calculation of correlations between the explanatory variables used in the analysis, and the actual k-means cluster algorithm, which has been conducted in three steps. In the first step, the idea is to derive a meaningful number of clusters by using a comprehensive dataset covering the period from 1995 to 2020.

In the second step, the dataset will be enlarged by including the forecast for all explanatory variables for the period between 1995 and 2025. As observed from Table 1, the forecasts for future trends of the aviation operational related variables are obtained from the informal interviews that were conducted among the experts from the airline industry, as well as from available resources, such as IATA reports.

In the third step, key measures from the PCA analysis will be interpreted to observe whether there are some structural changes between the previous crises and the COVID-19 outbreak, explore similarities/differences between the clusters over the time, and to check how the industry will recover. This can be achieved using the loading scores of the explanatory variables, which represent the linear combination for each principal component, i.e. the direction in which the new axes are pointing, derived during the PCA process. In addition, the contribution of each independent variable to the corresponding principal component provides information on the driving operational and financial factors within the clusters. Knowledge of both, the loading scores, and the contributions, will allow the years within the time series to be transformed from one cluster to another. However, the deeper understanding of the driving factors gained in this way must then be converted back into operating profit, leading to the application of system dynamics modeling, which is explained in the next Section 3.4.

3.4. System dynamic modeling

The decision to use a model to simulate and predict an airline's operating profit based on system dynamics was made for several reasons, the most important being that this technique has been already widely used and well proven. An analysis in the form of a preliminary study showed that the application of simpler models, e.g. linear models, is not possible because of the highly correlated nature of the data and the strongly non-linear behavior. The application of black box models or related techniques from the field of machine learning was not applicable in the given time frame of the recorded data due to the limited amount of training and test data. In contrast to existing airline operating profit models, the new model focuses on the approach of including carefully selected exogenous variables and ensuring relatively straightforward predictability for the future.

The system dynamic model is used in combination with k-means clustering for two different purposes. First, the model makes predictions for the operating profit based on highly correlated and dynamic exogenous variables and provides input data for k-means. Second, the model described is used to relate the knowledge gained about the "pseudo steady State of the System" and the underlying explanatory variables to the performance measure of operating profit.

The central variable of the model is operating profit, i.e., EBIT (earnings before interest and taxes), which describes the difference between operating revenue and operating expenses. The revenue part is depicted at the left side of Fig. 2, while the expense part is shown on the right side. Operating revenue is the product of yield (revenue per revenue passenger miles) and RPM. Ancillary Yield, which is an important part in calculating operating revenue, consists of fees for checked baggage, in-flight catering, and other services generated beyond the sale of tickets (O'Connell and Warnock-Smith, 2013). It is modeled as one of the derived ratios already explained in Section 3.1 and contains the ratio of ancillary yield to total yield per PAX. The advantage of using the relative number instead of the absolute value is that the value showed a nearly constant behavior during the period from 1995 to 2003, followed by a transition period between 2004 and 2009 and a return to a constant behavior by 2019 at a level 20 percent higher than before, which can be seen in Fig. 3. Hence, the constant behavior from 2010 to 2019 is the reason why the variable is held constant for the forecast until 2025. The transition from the lower to the higher level is explained by the change in business models induced by the growth of low-cost carriers (see e.g., Bejar, 2009), which were first to recognize the importance of Internet as

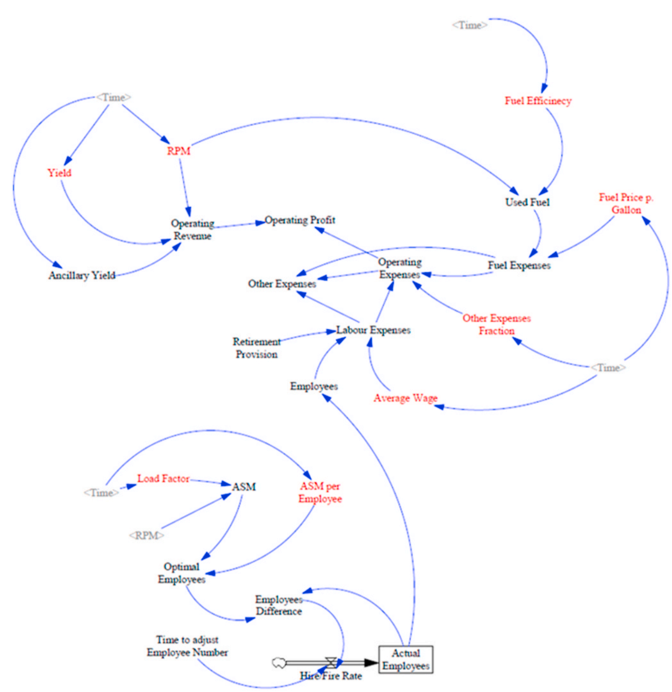


Fig. 2. The figure shows the system dynamics model used. The variables in red mark the exogenous variables. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

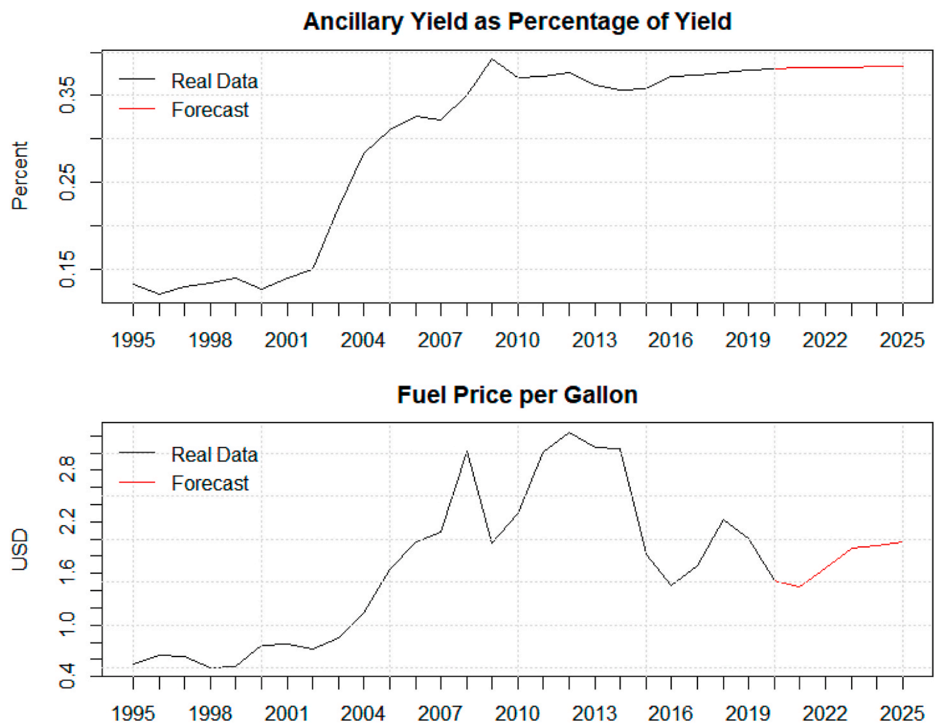


Fig. 3. The figure shows the development of ancillary yield as percentage of yield for the real and the forecasted data in the upper graph. The lower graph contains fuel price per gallon which is forecasted by using a regression model.

a facilitator of revenue generation.

Operating expenses, as the counterpart of operating revenue, is divided into fuel expenses, labor expenses and other expenses. The breakdown into these three types of expenses allows the impact of the crisis on the aviation industry to be modeled without committing to a detailed breakdown of costs, which would make forecasting difficult.

Fuel expenses are modeled by considering the price of fuel per gallon by RPM and an efficiency variable that describes the flown seat miles per gallon. Both ratios are derived from the variables listed in Table 1. The calculated metric of flown seat miles per gallon results from the division of total gallons of fuel and RPM, which showed a linear relationship with an increasing trend in the reviewed time period. This finding can be supported by the recent White paper published by ICCT White Paper (2019) which showed that fuel efficiency, measured in RPMs per gallon of fuel, improved by 3% from 2016 to 2018. The established trend was predicted as being ongoing until 2025 describing an improved fuel efficiency due to operational fuel optimization and increased technical efficiency. The fuel price and its forecast resulting from a regression model is already covered in Section 3.1.

Labor expenses are based on simulating the number of employees and their average wage. The recent study of Sobieralski (2020) suggested that total airline employment in the U.S will.

Decrease by 7%–13% with unskilled airline employees being the most severely impacted during these workforce reductions. However, it does not necessarily imply the reduction in average wage and thus, in overall labor expenses, as highly skilled and managerial positions would remain unaffected. The development of average wage after the crisis is first assumed to remain at the level of 2019 until the year 2023. After this period, an increase with a trend modeled on the four years before the COVID-19 crisis is assumed. This trend allows to roughly offset the annual inflation rate. The simulation of the total number of employees requires the use of an integral that describes the layoff and hiring of employees. This rate is the difference between the optimal and actual number of employees, where the optimal number of employees is calculated based on the available seat miles (ASM) and the ASM one employee can produce (ASM per employee). ASM itself is in turn dependent on RPM and the average load factor (LF). As already stated in Section 3.1, the data so far mainly comes from the MIT Airline Data Project (MIT, 2020). While the ASM produced per employee experienced a sharp increase between 2001 and 2005, a stagnant trend has been observed in recent years. Therefore, for the years beyond 2019 through 2025, produced ASM per employee is assumed to remain constant at 2019 levels, unaffected by the COVID-19 crisis. The load factor, which is a key metric for calculating Available Seat Miles (ASM) and optimal employees, drops from 85% to 59% in the U.S. market in 2020, according to IATA Economic Performance of the Airline Industry (IATA, 2020). Airlines are even more forced to increase load factors as quickly as possible once no bankruptcy protection is paid by governments. Therefore, it is assumed that the load factor will be raised to 70% in 2021. The load factor is forecasted to reach the pre-COVID-19 level of 85% in 2022 and follow a linear trend based on the years between 2012 and 2019 through the end of the forecast period in 2025. The rate of layoffs and hiring of new employees is adjustable by a constant called the “Time to adjust Employee Number”, which can be used to increase or decrease the rate to ensure faster layoff or hiring of staff. Parameter tuning has shown that a rate that is twice what it would normally be if simply taking the rate above, gives the best results.

The final part in the calculation of operating costs includes other expenses. These are also modeled as a derived metric as a percentage of total operating expenses. Data analysis of the years from 1995 to 2019 showed that other expenses accounted, on average, for 49% of total operating expenses and 50.1% on median (MIT, 2020). Since the variance is at a very low level of 0.002, it is assumed that the derived metric will be at 50.1% in the years from 2020 to 2025.

The next Section 4 now contains the results of both the system dynamic model, i.e. the airline operating profit, and the clustering analysis using k-means and PCA.

4. Results

After applying the methodology explained earlier, the system dynamic model provides the result shown in Fig. 4, where real and

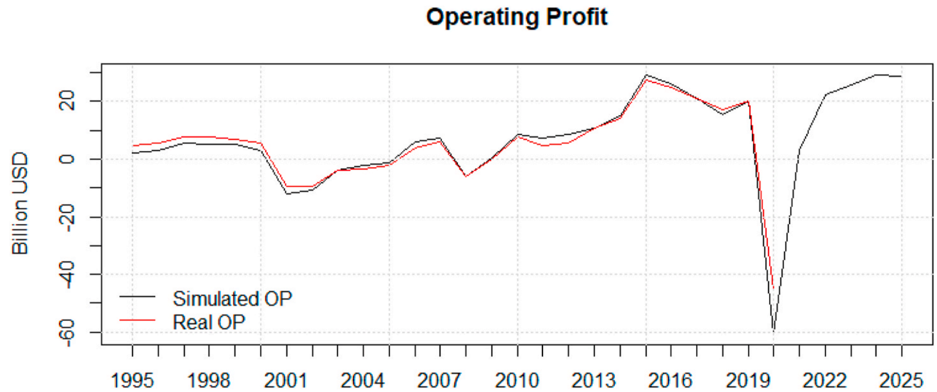


Fig. 4. Real and simulated operating profit.

modeled operating profit are compared. The system dynamic model can track the real operating profit in the years from 1995 to 2019 with high accuracy. During the COVID-19 induced collapse of the system in 2020 the operating profit was reported to be at −45 Billion USD where the model overshoots this value predicting an outcome of −60 Billion USD. Nevertheless, the model clearly shows that it can model such an enormous exogenous shock to a sufficient extent. The following years up to 2025 follow the forecasts in Table 1 based on literature and interviews, which include a relatively high degree of uncertainty in predicting the recovery of the aviation industry, as shown by the IATA report of April forecasting the recovery of revenue passenger kilometers (IATA, 2021). The scenario modeled in this way can be considered as a medium recovery scenario with respect to the IATA report.

The results of the k-means clustering methodology applied on the time span from 1995 to 2020 and on the time span from 1995 to 2025 including the recovery forecasts are given in Fig. 5 and Fig. 6 respectively. Both figures show a two-dimensional plot of PC1 and PC2. However, the interpretation of the result later, also considers at least 3 or even 4 PCs, which in the case of the 1995–2020 data account for 91% and 99% of the variation respectively. The importance of considering higher dimensional PCs is evident in Fig. 5, where 2014 belongs to cluster 5 but appears to be much closer to the center of cluster 2. This is no longer true when PC3 is added as an additional third dimension, where cluster 2 and 5 are clearly separated. However, this issue strongly demonstrates the strength of k-means where a simple two-dimensional plot can show a seven-dimensional clustering.

The clusters can be further analyzed by looking at the loading scores, i.e the linear combinations of the explanatory variables to the corresponding PC, and the contributions of each explanatory variable to the generated PC. The whole set of loading scores and contributions for PC1 until PC7 of each explanatory variable can be found in the Annex Table 5. These values are the output of the underlying PCA, which is needed for deeper understanding and interpretation of the two-dimensional plots.

Summarizing Annex Table 5, it needs to be pointed out that PC1 depends equally on all explanatory variables, which is represented by the similar amount of contribution of these variables to PC1. The reason for these equal contributions come from the relatively high correlation between many explanatory variables as visualized in Fig. 7. Moreover, Fig. 7 justifies the use of PCA as an appropriate method for dealing with

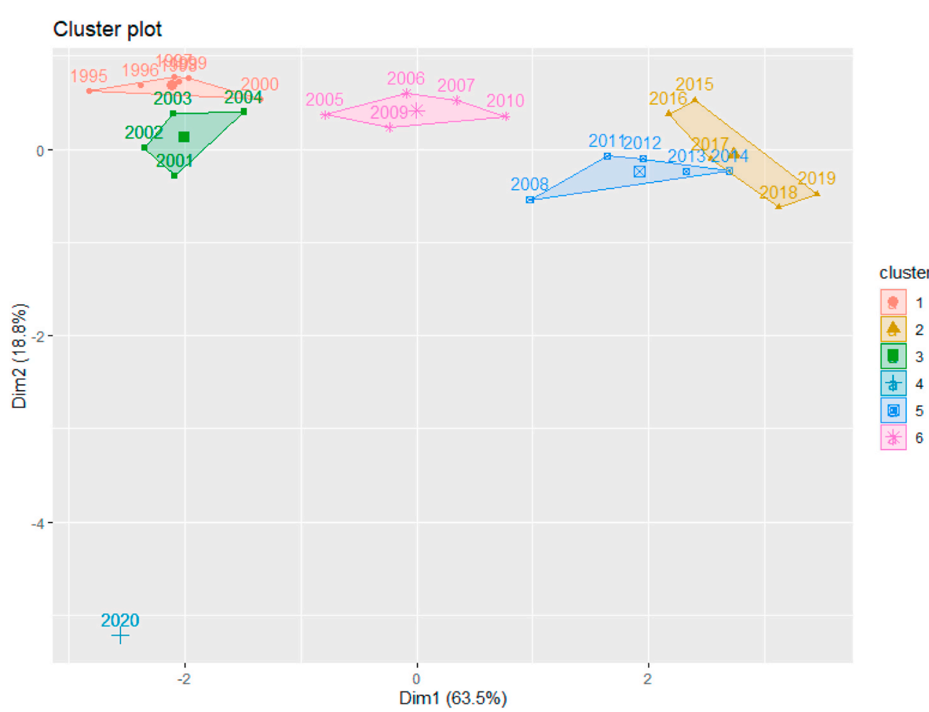


Fig. 5. Cluster plot for the time span of 1995–2020. The optimal number of clusters is 6. The year of 2020 where COVID-19 pandemic hit the airline industry is represented as a single cluster.

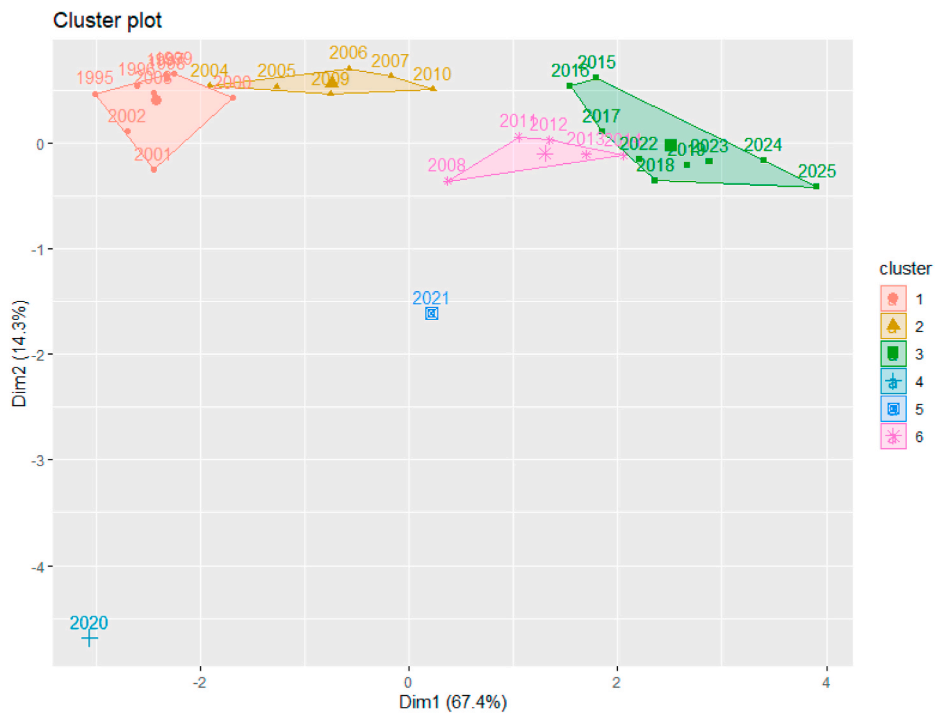


Fig. 6. Cluster plot for the time span of 1995–2025. The years 2020 and 2021 clearly represent two outliers within the rest of the simulated time span.

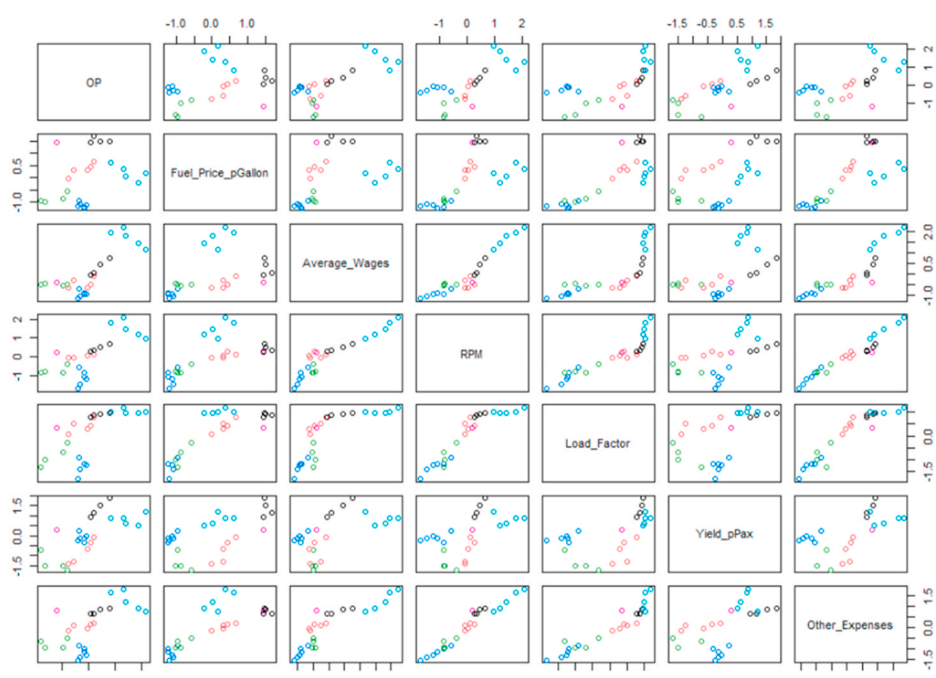


Fig. 7. Correlation plot of the used explanatory variables where the color stands for data points assigned to the same cluster. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

highly correlated data. Basically, the PCA method attempts to fit new axes to existing data in the direction that maximizes variance and separates the data better than before, i.e. with lower correlation. The usual approach of neglecting data that has a high correlation does not work in this example, as the application of the system dynamics model has shown that the choice of explanatory variables is important for the accuracy of the operating profit modeling.

The before mentioned equal contribution decreases with higher dimensional PC. Thus, operating profit and average wages already contribute 65% to PC2 and the inclusion of load factor results in a contribution of 80%. A focus on the loading scores for PC2, which describe the eigenvector indicate that the two major variables, operating profit and average wages, contradict each other in terms of direction of action. When considering PC3, which was often used for more sophisticated understanding in three dimensional plots, fuel price per gallon might be mentioned as by far the biggest driver with a contribution of over 65%.

The effect of high correlations between the explanatory variables and its summaries in the newly created PC1 leads to the interesting observation in Figs. 5 and 6, where an almost chronological sorting of the years by PC1 is visible. Both graphs indicate that the years from 1995 to 2020 and 2025, respectively, increase from left to right, with slight shifts in between. Considering the loading scores of PC1 which are all positive and close to one, result in an eigenvector which illustrates the growing trend in chronological order from left to right. The two outliers 2020 and 2021 do not fit this observation at first glance, but if we look at the difference between 2020 and 2021, we see that 2021 tends to slowly return to the pattern of the pre-COVID-19 era. The year 2020 is not only interesting due to the introduction of a single cluster far apart from the others, but it also introduces the need for considering PC4 for further analysis to keep the explained variations above 95% margin. Without including the COVID-19 crisis (years from 1995 to 2019), the explained variation remains at 97%. If 2020 is included in the PCA, the resulting explained variation decreases to 91% and increases to 99% when PC4 is added. PC4 now makes the same contribution to the total explained variation as PC3. In relation to all other explanatory variables yield per PAX is the clear driver of PC4, and hence the variable making 2020 unique to the years before.

Hence, it can be assumed that in 2020 the system was forced into a different state by the exogenous shock of the COVID-19 crisis but once the effect of the exogenous shock subsides, the system recovers, represented by the year 2021 as a transition cluster, and starts following old clusters again. This is illustrated in Fig. 6 where the years from 2022 up to 2025 cluster with pre-COVID years but also in the contributions of PC3 and PC4 to the overall explained variation in the data. There, the explanatory power of PC4 starts to decrease and the one of PC3 increases again, showing similar contributions as in the years before COVID-19.

For further analysis, the biplots (Fig. 8 to Fig. 10) are applied as an appealing tool in interpreting and visualizing the data after the application of the PCA. In addition to the individual data points, the variable and its correlations are also visualized. The color gradient of each vector shows the contribution in percent of the variable to the PC. For Fig. 8 it is the contribution of the variables to PC2 and for Fig. 9 and Fig. 10 it is the contribution to PC3. The contributions to PC1 are neglected in these diagrams, as it has already become clear that the contributions to PC1 are evenly distributed across all variables. Vectors pointing in the same direction with a small angle in between indicate a high positive correlation between each other, whereas arrows pointing in the opposite direction have a high negative correlation. Vectors perpendicular to each other are completely uncorrelated variables. In this vein, the cosine of the angle between the vectors describes the degree of correlation between the corresponding variables.

As observed from Fig. 8, it is noticeable that all vectors point in a

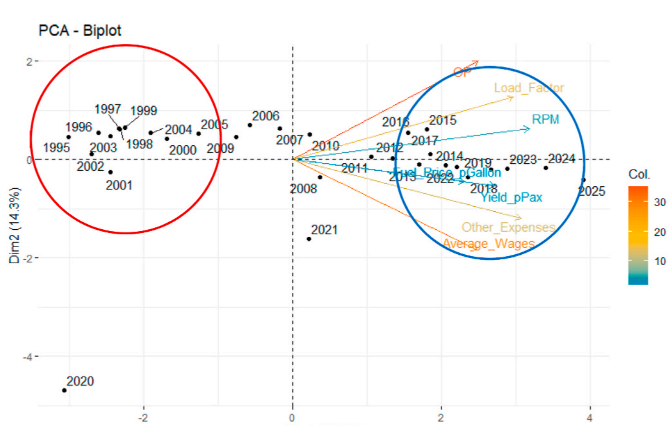


Fig. 8. Biplot of PC1 and PC2 including the contribution in percent of each variable to PC2 by the color gradient. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

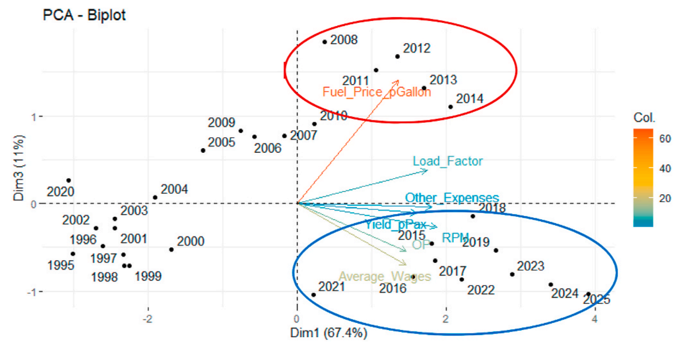


Fig. 9. Biplot of PC1 and PC3 including the contribution in percent of each variable to PC3 by the color gradient. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

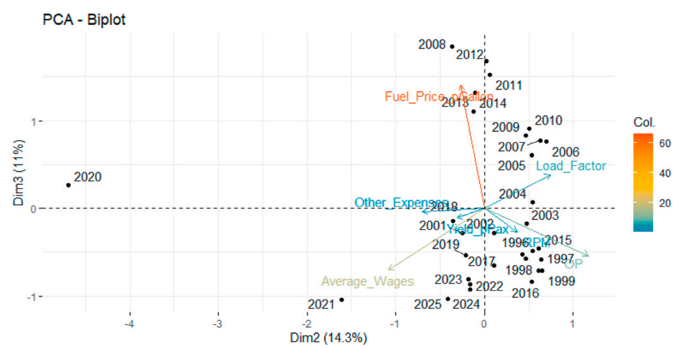


Fig. 10. Biplot of PC2 and PC3 including the contribution in percent of each variable to PC3 by the color gradient.

similar direction and thus, showing on the right side (blue circles) data points with relatively high values of these variables and on the left side (red circles) locating the data points with relatively low values. The small angles between many variables again illustrate what was said earlier about the high correlation between them. The fact that all vectors point in the same direction was to be expected as all loading scores for PC1 are positive and close to one, showing a clear growth trend. This increasing trend is then shown in the already mentioned effect of chronological order from 1995 on the left side and 2025 on the right side. External shocks such as the 9/11 or COVID-19 cause these points to move backward on the PC1 scale. However, the two most interesting vectors in Fig. 8 are represented by operating profit and average wages, which have a low correlation due to the opposite direction of effect shown in their loading scores for PC2. Moreover, these two variables make by far the biggest contribution to explain PC2. By considering load factor and to some extent other expenses, PC2 will be almost fully explained. Concluding the observation of strong correlation between variables in the positive direction of PC1, the hypothesis can be made that under the assumption of steadily increasing variables, future years will never be classified into previous clusters.

Fig. 9 now shows what was invisible in Fig. 5 and this includes PC3 as a third dimension, which is mainly explained by the fuel price per gallon which can be seen by the color gradient representing the contribution to PC3. In Fig. 5 two clusters were overlapping questioning the k-means algorithm doing a good job. However, now it gets clear that these clusters are well separated by PC3 which is illustrated with the red and blue circle.

Fig. 10, then shows what is to be expected when removing PC1 from the plot and showing PC2 to PC3 instead. Since PCA replaces the axes to better fit the data and thus most correlations are mapped to PC1, the correlation with higher dimensional PC decreases rapidly, which is visible in the direction of the vectors in Fig. 10. In contrast to the figures before the explanatory power of this figure is rapidly reduced and not very helpful for further considerations. The most interesting observation concerns the fuel price per gallon, where years with a high fuel price differ from those with a lower fuel price.

To answer the question of how the clusters differ from each other and which are similar, PCA by cluster was performed. Table 4 shows for each cluster how the explanatory variables will relate to PC1 and PC2. The color indicates the contribution of each explanatory variable to PC2. The contribution to PC1 is not visualized because the explanatory variables have a much higher correlation in relation to PC1 and thus have much more equal contributions, i.e. most of the explanatory variables have an important impact on PC1. It is immediately apparent from Table 4 that cluster "9/11" is most different from all the others in two respects. First, the correlations between the variables are completely different from those in the other clusters, and second, the contributions are much more evenly distributed with respect to PC2. This suggests that a third PC should be considered when trying to adequately represent and explain cluster "9/11" and its variation. The same is true for cluster "pre-& post-COVID", where a third dimension is needed to explain about 93% of the total variation in the data set, which is then comparable to the other two clusters. Clusters "Financial Crisis" and "post Financial Crisis" are similar in that the correlations of most variables are comparable and the explained variation within the data is close to 93% with only two dimensions. The biggest difference is in the way PC2 is represented, which is mainly driven by average wages in cluster "Financial Crisis", and other expenses as well as fuel price per Gallon in cluster "post Financial Crisis". The years of 2020 and 2021 cannot be visualized in the same way as they represent single clusters, but from analysis before it is known that the year 2020 requires the introduction of a new PC which is mostly driven by yield per PAX. Table 3 summarizes the above made observations by marking the explanatory variables in green, which are recognized as drivers for the corresponding PCs.

The knowledge gained from k-means clustering, which allows to represent the seven-dimensional clusters on an only two-dimensional plot and the underlying PCA for more advanced understanding of how the explanatory variables contribute to the new defined PCs can now be used for better understanding of future and past behavior of the system. In addition, the system dynamic approach must be used to transform the knowledge gained about the explanatory variables back into the airline profit cycles. The next Section 5 summarizes the most important aspects of the research conducted.

5. Conclusion

Understanding of the cycles in the airline industry is an imperative for airline managers and other stakeholders. The cyclical behavior of the airline industry imposes substantial challenges for analysts as it is affected by a great number of interrelated factors. The paper proposed a novel methodology combining two approaches to examine the airline profit cycle. On the one hand, the k-means/PCA methodology is applied to derive the meaningful number of clusters of years that are most similar with respect to the set of explanatory variables, reflecting the financial and operational performance of the airline industry. On the other hand, the system dynamic approach, which enables better understanding of a dynamically complex environment, is employed to forecast operating profit for the usage in the k-means methodology and to transform the gained knowledge from k-means methodology back into airline operating profit. Among a plethora of exogenous factors that hit the airline industry in the past, it appears that the COVID-19 pandemic has had an unprecedented impact on airlines' traditionally slim profitability. This study tackled the recovery pathways of airline profit cycles following the COVID-19 pandemic and focused on North America's air transport market.

The simulated pre-COVID-19 and COVID-19 era from 1995 to 2020 was supplemented by a post-COVID-19 forecast until 2025. The post-COVID-19 era is based on forecasts for the exogenous variables relying on available reports and expertise, describing a medium recovery scenario in terms of RPM. The forecasted operating profit is reasonable enough, given the satisfactory high accuracy of the model is achieved for the years prior to COVID-19. Nevertheless, the forecast contains a notable amount of uncertainty, as the development of passenger demand and the course of the COVID-19 pandemic is very hard to predict.

The k-means/PCA approach yielded the six different clusters for the period from 1995 to 2025. The first principal component reflects the continuous upward trend of most variables by placing them in chronological order. PC2 can be described as driven by operating profit, average wage and load factor and PC3's only driver is fuel price per gallon. The key measures from PCA, such as the equal contribution of each variable to PC1 or the consistently positive loading values close to 1, imply that under the assumption of continuous growth and a non-exogenous shock, future years will not cluster in past years. Furthermore, the analysis revealed that through the exogenous shock of COVID-19 a further dimension, mainly driven by yield per PAX, must be introduced. This explanatory power of the new dimension decreases to the end of the forecasting period, which raises the conclusion that post-COVID-19 years will follow the years of the pre-COVID-19 era. Hence, the "pseudo–Steady State of the System" which includes pre-COVID-19 and post-COVID-19 years can be explained by the underlying PCA measures. The method additionally shows how the system stays in a certain state for a few years and then drifts further to a new state. There are only a few variables that change to transfer from one cluster to the next. This knowledge of the transition to a new internal state increases the predictive power of future forecasts.

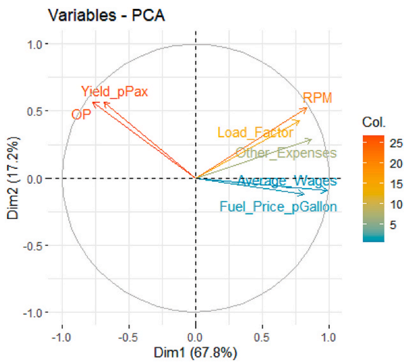
While this study provides a valuable understanding into the potential behavior of airline profitability cycles, research is already underway into how not only the magnitude of the explanatory variables affects the profit cycle, but also what impact a phase shift between the explanatory variables has on the results. This would allow us to convert points from one cluster to another and begin to better manage operating profit

Table 3
The explanatory variables marked in green can be considered as the driving variables due to their high contribution compared to the other variables in the corresponding cluster and PC. The years contained in the four clusters are linked in *Table 4*.

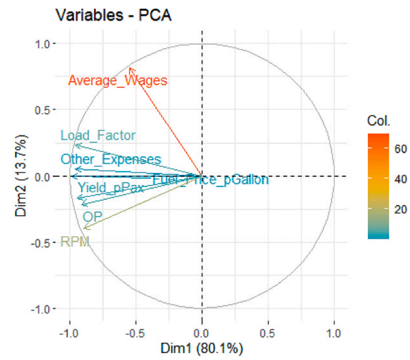
	Cluster "9/11"			Cluster "Financial Crisis"		Cluster "post Financial Crisis		Cluster "pre-& post-COVID"		
	PC1	PC2	PC3	PC1	PC2	PC1	PC2	PC1	PC2	PC3
Average Wage	■				■	■		■		■
Fuel Price p. Gallon	■			■			■		■	■
Load Factor	■	■	■	■		■		■		
Operating Profit	■	■	■	■		■			■	
Other Expenses	■		■	■			■	■		
RPM	■	■		■		■		■		
Yield p. PAX	■	■	■	■		■		■		■

Table 4
Output of the cluster wise PCA showing the correlation plot of each cluster. The color gradient indicates the contribution of the individual explanatory variables to PC2.

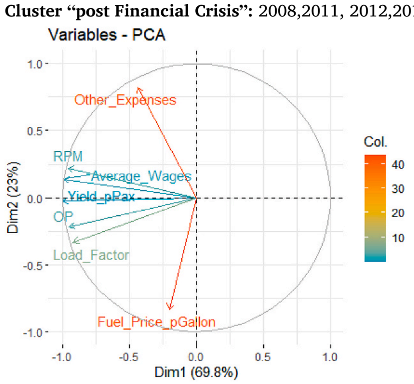
Cluster "9/11": 1995,1996,1997,1998, 1999,2000, 2001,2002,2003



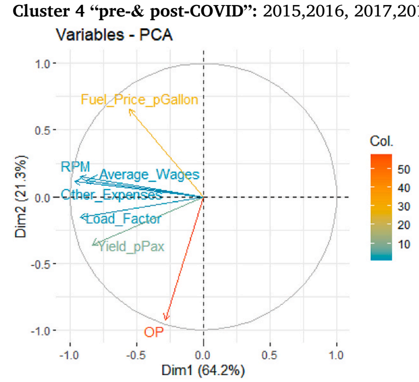
Cluster "Financial Crisis": 2004,2005,2006, 2007,2009,2010



Cluster "post Financial Crisis": 2008,2011, 2012,2013,2014



Cluster 4 "pre-& post-COVID": 2015,2016, 2017,2018,2019,2022, 2023,2024,2025



Cluster "COVID": 2020, 2021
PCA not feasible

through the influence of its explanatory variables.

Finally, as seen from the previous disruptive events, the recovery pathway may substantially differ for airlines with different business models. The results of a recent study conducted by Kaffash and Khezrimotlagh (2022) demonstrated that U.S. low-cost carriers have a higher efficiency than network carriers in pandemic. In addition, the study also claimed that U.S. network carriers respond more aggressively to COVID-19 pandemic in terms of flight cancellations, reduction in scheduled flights. Similar to the post-2001 crisis, it seems that LCCs are able to overcome the crisis in a more efficient way, probably attributed to their lower operating costs, lower concentration on international routes and the market segment served. Since the effects of pandemic are still present, another interesting future study could investigate how the potential redistribution of market share between LCCs and FSCs may affect the behavior of airline profitability. However, the model *per se* is not able to capture the effect of potential change of market share between FSCs and LCCs without further modification of the underlying model mechanism. The main reason resides in the fact that the model is designed to predict the U.S airline industry as a whole, at its most aggregate level. Thus, we leave these adaptations for future research.

Author statement

Manuel Renold: Conceptualization, Methodology, Validation, Investigation, Writing – original draft, Visualization, Janik Vollenweider: Investigation, Validation, Writing – original draft, Visualization, Data curation, Formal analysis, Software, Nemanja Mijović; Investigation, Validation, Writing – original draft, Visualization, Data curation, Formal analysis, Software, Jovana Kuljanin: Project administration, Investigation, Writing – original draft, Formal analysis, Validation, Supervision, Milica Kalić: Validation, Writing – review & editing. Herein, we would like to emphasize that Janik Vollenweider, Nemanja Mijović and Jovana Kuljanin shared the equal effort during the preparation of the manuscript. Thus, they can be considered as a second author.

Appendix

Fig. 10Development of the set of explanatory/input variables used in the system dynamic model and for the PCA analysis.

Table 4
Exogenous input variable for the airline operating profit model

Year	Operating Profit	Fuel Price p. Gallon	Average Wages	Seat Miles p. Gallon	Ancillary Yield	ASM p. Employee	Other Expenses	Load Factor	Yield p. Pax	RPM
1995	2347800000	0.55	47486	36.45	0.13	1902087	2.06	0.67	0.13	486535059000
1996	3131060000	0.65	49128	37.65	0.12	1968180	2.03	0.70	0.13	519374038000
1997	5571280000	0.63	50337	38.19	0.13	1949526	2.03	0.71	0.13	548424210000
1998	4971120000	0.50	51275	38.20	0.13	1898793	2.08	0.71	0.13	563980242000

(continued on next page)

Table 4 (*continued*)

Year	Operating Profit	Fuel Price p. Gallon	Average Wages	Seat Miles p. Gallon	Ancillary Yield	ASM p. Employee	Other Expenses	Load Factor	Yield p. Pax	RPM
1999	4972890000	0.52	50814	38.77	0.14	1876632	2.07	0.71	0.13	594551011000
2000	2897040000	0.76	54243	38.83	0.13	1887786	1.94	0.73	0.13	629665842000
2001	−12169600000	0.78	57286	39.23	0.14	1881177	1.96	0.70	0.12	594871149000
2002	−10572700000	0.71	58371	42.11	0.15	1949241	1.91	0.72	0.11	599548717635
2003	−3713210000	0.85	56664	44.71	0.22	2084240	1.96	0.74	0.11	596247765924
2004	−2000290000	1.13	57138	46.02	0.28	2285684	2.02	0.76	0.11	656245961436
2005	−1461140000	1.64	55358	48.52	0.31	2489519	2.02	0.79	0.11	693619850711
2006	6088250000	1.97	54852	50.29	0.33	2569977	1.99	0.80	0.12	710627624916
2007	7354790000	2.08	57394	50.99	0.32	2596364	1.97	0.81	0.13	737360156433
2008	−6074310000	3.02	58821	51.87	0.35	2608664	1.91	0.80	0.13	729139953962
2009	621383000	1.95	60550	52.84	0.39	2548903	2.02	0.81	0.12	692570116535
2010	8727430000	2.29	63574	54.37	0.37	2603850	1.95	0.83	0.13	717394758419
2011	7305310000	3.01	64058	54.99	0.37	2615944	1.89	0.83	0.14	736688612526
2012	8677130000	3.24	65981	56.62	0.38	2605362	1.84	0.83	0.14	747651705646
2013	10922200000	3.07	71899	57.21	0.36	2715112	1.87	0.84	0.15	765044741853
2014	15024600000	3.05	76463	57.35	0.36	2741862	1.85	0.84	0.15	786728156369
2015	29072800000	1.82	82741	58.23	0.36	2753700	1.96	0.84	0.14	826405762444
2016	26267400000	1.46	87175	58.39	0.37	2727903	2.00	0.84	0.14	855916670087
2017	20922300000	1.69	92379	59.49	0.37	2749000	1.94	0.84	0.14	887311636743
2018	15275800000	2.23	94764	60.70	0.38	2815807	1.89	0.84	0.14	930745213031
2019	20263800000	2.01	98927	61.48	0.38	2821188	1.90	0.85	0.14	969997168956
2020	−60660600000	1.52	102345	44.00	0.65	1790000	2.03	0.59	0.13	339648000000
Forecast Data										
2021	3055000000	1.45	98927	63.12	0.38	2821188	1.96	0.70	0.14	705213239357
2022	22328300000	1.66	98927	63.92	0.38	2821188	1.96	0.85	0.14	857871164112
2023	25835900000	1.89	98927	64.72	0.38	2821188	1.96	0.85	0.14	942478707673
2024	29376300000	1.93	102722	65.50	0.38	2821188	1.96	0.85	0.15	1008917912571
2025	28899700000	1.96	106486	66.28	0.38	2821188	1.96	0.86	0.15	1066930692544

Table 5
Loading scores from the PCA applied on different time spans and clusters

	OP	Fuel Price p. Gallon	Average Wages	RPM	Load Factor	Yield p. Pax	Other Expenses
Time Span from 1995 to 2020							
LoadingScores.Dim.1	0.641	0.784	0.650	0.919	0.902	0.745	0.887
LoadingScores.Dim.2	0.720	−0.208	−0.616	0.257	0.341	−0.082	−0.430
LoadingScores.Dim.3	−0.212	0.580	−0.435	−0.168	0.114	0.058	−0.032
LoadingScores.Dim.4	0.137	−0.055	−0.020	−0.207	−0.225	0.658	−0.145
LoadingScores.Dim.5	−0.072	−0.044	−0.086	0.127	−0.058	0.039	0.047
LoadingScores.Dim.6	−0.043	−0.002	0.024	0.027	0.037	0.017	−0.064
LoadingScores.Dim.7	0.015	0.029	0.013	0.022	−0.038	−0.011	−0.022
Contribution.Dim.1	9.225	13.824	9.487	19.004	18.301	12.470	17.688
Contribution.Dim.2	39.445	3.281	28.874	5.035	8.827	0.515	14.022
Contribution.Dim.3	7.263	54.566	30.774	4.575	2.110	0.548	0.163
Contribution.Dim.4	3.277	0.523	0.069	7.525	8.898	76.025	3.684
Contribution.Dim.5	13.629	5.066	19.507	43.011	8.844	4.101	5.843
Contribution.Dim.6	20.929	0.058	6.628	7.939	15.541	3.070	45.835
Contribution.Dim.7	6.232	22.681	4.662	12.910	37.479	3.272	12.765
Time Span from 1995 to 2025							
LoadingScores.Dim.1	0.741	0.684	0.737	0.945	0.881	0.809	0.912
LoadingScores.Dim.2	0.594	−0.134	−0.548	0.187	0.376	−0.157	−0.357
LoadingScores.Dim.3	−0.274	0.712	−0.357	−0.137	0.194	−0.056	−0.019
LoadingScores.Dim.4	0.126	0.042	−0.114	−0.192	−0.178	0.561	−0.168
LoadingScores.Dim.5	−0.047	−0.033	−0.124	0.114	−0.076	0.030	0.091
LoadingScores.Dim.6	−0.067	−0.055	−0.014	−0.010	0.086	0.040	−0.001
LoadingScores.Dim.7	0.027	−0.012	−0.021	−0.058	0.008	−0.008	0.063
Contribution.Dim.1	11.637	9.931	11.519	18.928	16.448	13.895	17.642
Contribution.Dim.2	35.296	1.809	30.033	3.511	14.158	2.471	12.721
Contribution.Dim.3	9.771	65.876	16.568	2.434	4.898	0.406	0.047
Contribution.Dim.4	3.570	0.407	2.939	8.367	7.196	71.138	6.384
Contribution.Dim.5	4.655	2.262	32.877	27.967	12.502	1.919	17.817
Contribution.Dim.6	26.606	18.050	1.164	0.603	44.105	9.465	0.008
Contribution.Dim.7	8.464	1.664	4.900	38.189	0.694	0.707	45.382

References

Adler, N., Berechman, J., 2001. Measuring airport quality from the airlines' viewpoint: an application of data envelopment analysis. Transport Pol. 8 (3), 171–181.

AGM, 2020. Annual general meeting, outlook for air transport, and the airline industry – november 2020. Available at: https://www.iata.org/en/events/agm/annual-general-meeting-2020/. (Accessed 12 March 2021).

Alamdari, F., Fagan, S., 2005. Impact of the adherence to the original low-cost model on the profitability of low-cost airlines. Transport Rev. 25 (3), 377–392.

Barnhart, C., Farahat, A., Lohatepanont, M., 2009. Airline fleet assignment with enhanced revenue modeling. Oper. Res. 57 (1), 231–244.
Bejar, R., 2009. Airline trends and ancillary revenue report 2010. Available at. http://www.airsavings.net/whitepapers/Airsavings%202010%20ForecastJan2010.pdf. (Accessed 27 May 2021).
Button, K., McDougall, G., 2006. Institutional and structure changes in air navigation service-providing organizations. J. Air Transport. Manag. 12 (5), 236–252.
Chin, A.T., Tay, J.H., 2001. Developments in air transport: implications on investment decisions, profitability and survival of Asian airlines. J. Air Transport. Manag. 7 (5), 319–330.
Choi, J.W., O'Connor, M., Truong, D., 2019. Predicting the U.S. Airline operating profitability using machine learning algorithms. International journal of aviation. Aeronautics and Aerospace 6 (5), 1–31.
Collins, D.L., Román, F.J., Chan, H.C., 2011. An empirical investigation of the relationship between profitability persistence and firms' choice of business model: evidence from the US airline industry. J. Manag. Account. Res. 23 (1), 37–70.
Cronrath, E.-M., 2018. The Airline Profit Cycles: A System Analysis of Airline of Airline Industry Dynamics. Routledge, New York.
Deloitte, 2020b. Price forecast Oil, gas & chemicals. Available at. https://www2.deloitte.com/ca/en/pages/resource-evaluation-and-advisory/articles/deloitte-canadian-price-forecast.html. (Accessed 21 March 2021).
Dobruszkes, F., Van Hamme, G., 2011. The impact of the current economic crisis on the geography of air traffic volumes: an empirical analysis. J. Transport Geogr. 19 (6), 1387–1398.
Doganis, R., 2005. Airline Business in the 21st Century. Routledge.
ICAO, 2020. Effects of Novel Coronavirus (Covid-19) on Civil Aviation: Economic Impact Analysis. Available at:. https://www.icao.int/sustainability/Documents/COVID-19/ICAO_Coronavirus_Econ_Impact.pdf. (Accessed 6 May 2021).
Elhmoud, E.R., Kutty, A.A., Abdalla, G.M., Kucukvar, M., Bulak, M.E., Elkharaz, J.M., 2021. Eco-efficiency performance of airlines in eastern Asia: a principal component analysis based sustainability assessment. In: In Proceedings of the 11th Annual International Conference on Industrial Engineering and Operations Management, Singapore, pp. 7–11.
EUROCONTROL, 2022. Forecast update 2022-2024: European flight movements and service units - recovery from COVID-19 and Russian invasion of Ukraine. Available at: https://www.eurocontrol.int/sites/default/files/2022-06/eurocontrol-three-year-forecast-2022-2024-june-2022.pdf. (Accessed 15 July 2022).
Franke, M., 2007. Innovation: the winning formula to regain profitability in aviation? J. Air Transport. Manag. 13, 23–30.
Franke, M., John, F., 2011. What comes next after recession? Airline industry scenarios and potential end games. J. Air Transport. Manag. 17, 19–26.
Hätty, H., Hollmeier, S., 2003. Airline strategy in the 2001/2002 crisis—the Lufthansa example. J. Air Transport. Manag. 9, 51–55.
IATA, 2020. Economic performance of the airline industry – 2020 end year report. Available at. https://www.iata.org/en/iata-repository/publications/economic-reports/airline-industry-economic-performance–november-2020–report/. (Accessed 29 December 2020).
IATA, 2021. Outlook for the global airline industry. – April 2021. Available at: https://www.iata.org/en/iata-repository/publications/economic-reports/airline-industry-economic-performance–april-2021–report/. (Accessed 13 September 2021).
IATA Economics, 2013. Profitability and the air transport value chain. Retrieved from. https://www.iata.org/en/iata-repository/publications/economic-reports/profitability-and-the-air-transport-value-chain/. (Accessed 17 May 2021).
ICCT (International Council on Clean Transportation) White Paper, 2019. U.S. Domestic Airline FuelEfficiency Ranking, 2017-2018. By Xinyi Sola Zheng; Brandon Graver, Ph.D.; Dan Rutherford, Ph.D.
Jiang, H.H., Hansman, R.J., 2004. An analysis of profit cycles in the airline industry. MIT International Center for Air Transportation, ICAT-2004-7.
Kaffash, S., Khezrimotlagh, D., 2022. U.S. network and low-cost carriers' performance in response to COVID-19: strictness of government policies and passengers' panic. Research in Transportation Business & Management. https://doi.org/10.1016/j.rtbm.2022.100835. In press.
Lau, R.S.M., Mattheiss, T.H., 1992. A Markov model of the growth and profitability of the US airline industry. Logist. Transport Rev. 28 (2), 189.
Lenoir, N., 1998. Cycles in the Air Transportation Industry, Paper Presented at the 8th World Conference on Transportation Research. July 1998. Belgium, Antwerp, pp. 12–17.
Liehr, M., Größler, A., Klein, M., Milling, P.M., 2001. Cycles in the sky: understanding and managing business cycles in the airline market. Syst. Dynam. Rev. 17 (4), 311–332.
Lyneis, J.M., 2000. System dynamics for market forecasting and structural analysis. Syst. Dynam. Rev.: The Journal of the System Dynamics Society 16 (1), 3–25.
Martini, G., 2022. The Air Transportation Vertical Channel, the Global Value Added, and the Role Played by Private versus Public Control. In: The Air Transportation Industry. Elsevier, pp. 77–97.
Maung, Y., Douglas, I., Tan, D., 2022. Identifying the drivers of profitable airline growth. Transport Pol. 115, 275–285.
Mijović, N., Kalić, M., Kuljanin, J., Renold, M., 2018. Airline profitability cycles: an undamped system model approach. In: Proceedings: XIII Balkan Conference on Operational Research Proceedings, pp. 398–405.
Miranda, U., 2015. The relationship between terrorism, oil prices, and airline profitability (Doctoral dissertation). In: Available from Walden Dissertations and Doctoral Studies Collection at ScholarWorks.
Mit, 2020. Airline data Project. Available at. http://web.mit.edu/airlinedata/www/default.html. (Accessed 10 September 2021).
O'Connell and Warnock-Smith, 2013. An investigation into traveler preferences and acceptance levels of airline ancillary revenues. J. Air Transport. Manag. 33, 12–21.
Pierson, K., Sterman, J.D., 2013. Cyclical dynamics of airline industry earnings. Syst. Dynam. Rev. 29 (3), 129–156.
Profillidis, V., Botzoris, G., 2015. Air passenger transport and economic activity. J. Air Transport. Manag. 49, 23–27.
Roucolle, C., Seregina, T., Urdanoz, M., 2020. Measuring the development of airline networks: comprehensive indicators. Transport. Res. Pol. Pract. 133, 303–324.
Shaked, I., OrelowItz, B., 2020. The airline industry and COVID-19: saving for a rainy day. American Bankruptcy Institute Journal 39 (5), 36–40.
Sobieralski, B.J., 2020. COVID-19 and airline employment: insights from historical uncertainty shocks to the industry. Transp. Res. Interdiscip. Perspect. 5.
Statista, 2020. Number of scheduled passengers boarded by the global airline industry from 2004 to 2020. Available at. https://www.statista.com/statistics/564717/airline-industry-passenger-traffic-globally/. (Accessed 15 July 2022).
Suau-Sanchez, P., Voltes-Dorta, A., Cuguero, 2020. An early assessment of the impact of COVID-19 on air transport: just another crisis or the end of aviation as we know it? J. Air Transport. Manag. 86.
Tarry, C., 2015. Discerning the GDP multiplier effect. Airl. Bus. 31 (7), 52–55.
The Economist, 2014. Why airlines make such meagre profits. Available at. http://www.economist.com/blogs/economist-explains/2014/02/economist-explains-5. (Accessed 16 July 2022).