

## Original Article

<https://doi.org/10.12985/ksaa.2023.31.1.099>  
ISSN 1225-9705(print) ISSN 2466-1791(online)

## Analyzing the Impact of Pandemics on Air Passenger and Cargo Demands in South Korea

Jungtae Song\*, Irena Yosephine\*\*, Sungchan Jun\*\*, Chulung Lee\*\*\*

### ABSTRACT

글로벌 팬데믹 사태는 항공 수요에 부정적인 영향을 끼치는 요소 중 하나다. 글로벌 팬데믹으로 인해 한국은 2020년과 2021년의 항공 승객 수가 2019년 대비 각각 68.1%와 47% 감소했다. 본 연구는 지난 20여년 동안 발생한 4대 팬데믹 특성을 분석, 전염병의 영향을 연구하는 것을 목표로 한다. SARS, H1N1, MERS 및 COVID-19의 발생기간 동안 한국의 항공 여객 및 화물 수요에 대한 실증 데이터를 활용하여 영향력을 분석한다. 또한 머신러닝 회귀 모델을 구축하여 향후 발생할 다른 전염병 대한 항공 수요를 예측하고자 한다. 연구 결과, 전염병이 항공 운항편수와 승객에 부정적인 영향을 미친다는 사실을 발견하였다. 반면화물 수송에는 긍정적인 영향을 미친다는 분석 결과를 도출하였다. 본 분석에 활용되는 회귀 모델은 팬데믹 기간 동안 항공 수요를 예측하는 데 평균 86.8%의 기능을 보였다. 또한 본 연구는 특정 국가의 팬데믹 상황보다 전 세계적인 팬데믹 상황이 항공 운송 수요에 더 많은 영향을 미친다는 것을 보여준다.

**Key Words** : Air Transportation Demand(항공 수요), Pandemic(팬데믹), Passenger(여객), Cargo(화물), Machine Learning(머신러닝), Regression Model(회귀 모델)

### 1. INTRODUCTION

Tourism is one of the most vulnerable sectors to global economic and financial crises, recession, terrorism, political shocks, natural disasters, epidemics, and pandemics (Collins-Kreiner and Ram, 2021). The tourism industry is a catalyst for the spread of the virus and its victims. Therefore, this sector is one of the hardest hit by the pandemic. The tourism industry has experienced several significant

pandemics over the past few decades—most notably, severe acute respiratory syndrome (SARS) (2002–2004), novel influenza A (H1N1) (2009–2010), MERS (2015), and COVID-19 (2019–present) (Hu et al., 2021). During the COVID-19 pandemic, airlines were the biggest destroyer of value among all aviation subsectors, estimated to suffer an economic loss of 167.9 billion USD in 2020 (McKinsey & Company, 2022). In South Korea, the number of air passengers in 2020 and 2021 declined by 68.1% and 47%, respectively, compared to 2019. Losses are expected to be 39.4 million and 6.38 million (in U.S. dollars) in 2020 and 2021, respectively (KOTI, 2021). Given the possibility that such a crisis will occur again in the future, various precautions, such as secured financial resources, must be taken (Park and Jeon, 2020).

However, the air cargo sector has suffered comparatively less than the passenger sector,

Received: 18. Jan. 2023, Revised: 16. Feb. 2023,

Accepted: 06. Mar. 2023

\* Research Professor, Hanseo University

\*\* Graduate Student, Industrial Management Engineering, Korea University

Corresponding Author E-mail :

irena.yosephine@gmail.com

Corresponding Author Address : 145 Anam-ro, Seongbuk-gu, Seoul, Korea

\*\*\* Department Head, Industrial Management Engineering, Korea University

and it was expected that it would recover before the passenger sector (Wang et al., 2021). Except for freight forwarders and cargo airlines, all segments of the aviation value chain reported massive losses in 2020. These sub-sectors benefited from a rise in demand for air cargo. In 2020, both managed to generate healthy economic profits: For freight forwarders, 4%; and for air cargo carriers, 9% (McKinsey & Company, 2022). Despite the rapid decline in the passenger sector, air cargo exports in South Korea increased by about 4.4% year on year (January - April 2020) (Kim et al., 2020).

Therefore, this study aims to study the impact of the pandemic on passenger and cargo demand, separately, in the aviation industry in South Korea. It uses empirical data from Incheon International Airport. Furthermore, understanding the impact of the pandemic will be a basis for predicting air transportation demand to anticipate other pandemics in the future.

Accordingly, the remainder of this study can be categorized as follows. Section 2 provides an overview of relevant research and existing literature. Section 3 introduces the regression model framework and the dataset that is being studied. Section 4 reports the computational results, including the correlation coefficient between inputs and outputs, regression model performance, feature importance, and results analysis. An overview of future research directions is provided in Section 5.

## II. LITERATURE REVIEW

### 2.1 Previous Studies on Analyzing Pandemic Impact on Passenger and Cargo Demand

The number of cases (infected patients) is used to assess the impacts of SARS and Avian Flu on international tourism demand to Asia, which showed the negative impact of SARS and the non-significant impact of Avian Flu on the number of foreign visitors (Kuo et al., 2008). In

addition, the number of deaths is also used to assess the impact of COVID-19 on transport volume and freight capacity dynamics of food retail logistics in Germany, which showed a strong linear statistical relationship between dry product transport volume growth and the number of new COVID-19 cases (Loske, 2020).

Other than the number of cases and a number of deaths, the World Pandemic Uncertainty Index (WPUI) and Discussion about Pandemic Index (DPI) created by Ahir et al. (2018) are used to assess the impact of pandemics on tourism, which showed the negative impact of pandemic uncertainty and discussion on tourist arrivals in the long run (Kocak et al., 2022). In the case of South Korea, the number of cases of SARS and MERS and the distance to South Korea also showed a negative impact of the pandemic on the number of arrivals (Joo et al., 2019).

This study uses data from Incheon International airport to assess the impact of pandemic-related features including the number of cases, number of deaths, WPUI, and DPI of four major pandemics (SARS, H1N1, MERS, and COVID-19) on airline demands for passenger and cargo transportation. A regression machine learning model is then developed to predict airline demand for air transportation.

### 2.2 Brief Information on Four Major Pandemics (SARS, H1N1, MERS, COVID-19)

The purpose of this paper is to assess the impact of pandemic-related features on passenger and cargo demand in the aviation industry by examining four major pandemics. Brief information about the pandemic is presented in Table 1.

Severe acute respiratory syndrome (SARS) is a viral respiratory disease caused by SARS-associated Coronavirus. SARS is an airborne virus and can spread through small droplets of saliva in a similar way to the cold and

Table 1. Brief information on pandemics

No	Pandemic	Duration	Total cases*	Total deaths*
1	SARS	9 months (Nov '02 – Jul '03)	8,096 (3)	774 (0)
2	H1N1	17 months (Apr '09 – Aug '10)	565,540 (5,922)	18,616 (201)
3	MERS**	3 months (May '15 – Jul '15)	1,384 (186)	495 (36)
4	COVID-19***	> 35 months (Dec '19 – Now)	574,940,809 (19,776,050)	6,403,638 (25,047)

\*World (South Korea), \*\*During outbreak in South Korea, \*\*\*As per July 2022.

Source: WHO and CDC South Korea website.

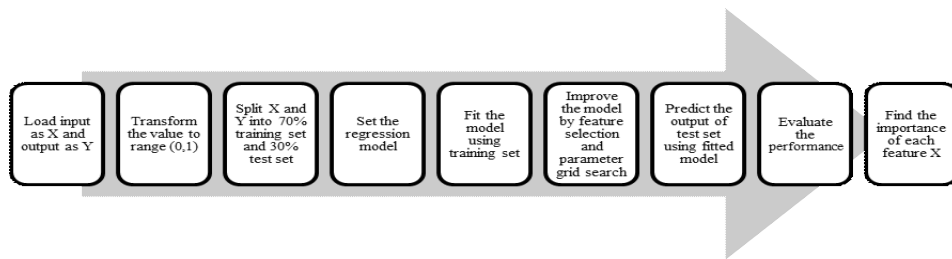


Fig. 1. Regression model framework

influenza. The case fatality rate for persons with illness meeting the current WHO case definition for probable and suspected cases of SARS is around 3%.

The influenza A (H1N1) virus showed that it was derived from an animal influenza virus. After initial reports of an influenza outbreak in North America in April 2009, a new influenza virus spread rapidly throughout the world. By the time WHO declared a pandemic in June 2009, a total of 74 countries and territories had reported laboratory-confirmed infections.

Middle East respiratory syndrome Coronavirus (MERS-CoV) is a virus that is transmitted to humans from infected dromedary camels. As a zoonotic virus, it can be transmitted between animals and humans through direct or indirect contact with infected animals.

Coronavirus disease (COVID-19) is an infectious disease caused by the SARS-CoV-2 virus. Coughing, sneezing, talking, singing, or breathing can spread viruses in tiny droplets. These particles range from large respiratory droplets

to smaller aerosols.

### III. METHODOLOGY

This study builds several machine learning regression models, such as decision trees, random forests, K-neighbors, and XG Boost (the choice of a regression model is not a major concern in this study), to forecast air transportation demand during pandemic periods. Feature selection and parameter grid search are used to improve the performance of the model and cross-validation (70% of the training set and 30% of the test set) prevent the model from overfitting. The model with the highest performance based on the *R*-squared value is chosen for further analysis. The detail of the regression model framework is shown in Fig. 1.

An empirical dataset from Incheon International Airport in South Korea with 257 monthly data from March 2001 to July 2022 is used. The inputs of the model are pandemic-associated features as defined in Table 2. The

Table 2. Pandemic-related features

Feature	Description
Pandemic_Timeframe	Categorical data 1: pandemic occur, 0: no pandemic
WPUI_World	World Pandemic Uncertainty Index
DPI_World	Discussion about Pandemics Index
WPUI_KOR	World Pandemic Uncertainty Index (South Korea only)
DPI_KOR	Discussion about Pandemics Index (South Korea only)
New_Cases_World	New (probable) infected case in the world per month
Cumulative_Cases_World	Cumulative (probable) infected case in the world per month
New_Deaths_World	New death case in the world per month
Cumulative_Deaths_World	Cumulative death case in the world per month
New_Cases_KOR	New (probable) infected case in South Korea per month
Cumulative_Cases_KOR	Cumulative (probable) infected case in South Korea per month
New_Deaths_KOR	New death case in South Korea per month
Cumulative_Deaths_KOR	Cumulative death case in South Korea per month

pandemic-related features are focused on the four major pandemics from 2001 to 2022: SARS, H1N1, MERS, and COVID-19. The output of the model is features related to airline demand at Incheon International Airport in South Korea based on empirical from this study. The air transportation demands are separated into three categories, which are the total number of flights (Flight Total), the total number of passengers (Passenger Total), and the total number of cargo (Cargo Total).

#### IV. RESULT ANALYSIS

In this paper, a regression model with the highest performance based on the  $R$ -squared value will be chosen for further analysis to see the impact on feature selection. Feature selection will depend on correlation analysis between inputs and outputs. Furthermore, we will conduct regression analyses specific to airline demands such as domestic and international departures and arrivals.

##### 4.1 Correlation Analysis between Inputs and Outputs

Based on the result shown in Fig. 2, all features of input have negative correlation with number of flight and number of passengers where overall have more negative correlation with number of passengers compared to number of flights, meanwhile has positive correlation with number of cargos.

The number of cases or deaths around the world has a higher correlation than the number of cases or deaths in only South Korea. Therefore, the number of cases or deaths in South Korea will be excluded from the regression analysis.

##### 4.2 Regression Model Performance

Based on the average  $R$ -squared value shown in Table 3, the XGBoost Regressor model gives the most accurate performance. Generally, the number of flights and passengers is similar,

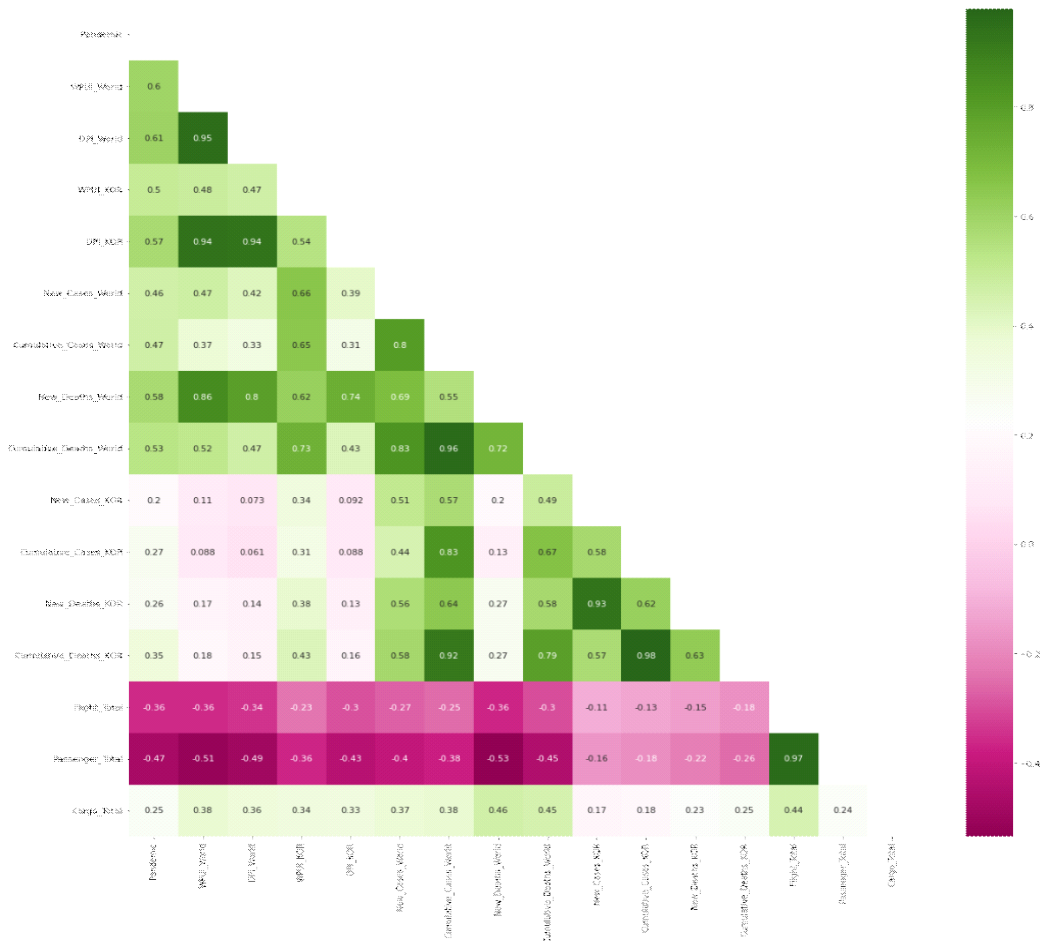


Fig. 2. Heatmap of correlation coefficient

indicating a positive correlation between the two outputs as indicated on the heatmap.

### 4.3 Feature Importance

Feature selection is conducted by excluding

Table 3.  $R$ -squared value for each regressor model

No	Model	Flight	Passenger	Cargo	Average
1	Decision tree	0.794	0.837	0.620	0.750
2	Random forest	0.787	0.764	0.689	0.747
3	K-Neighbour	0.769	0.776	0.620	0.722
4	XG Boost	0.934	0.920	0.751	0.868
Average		0.821	0.824	0.670	

the pandemic-related features that have a lower correlation coefficient value, which is the number of cases or deaths in South Korea only (New\_Cases\_KOR, Cumulative\_Cases\_KOR, New\_Deaths\_KOR, Cumulative\_Deaths\_KOR). According to the average  $R$ -squared value shown in Table 4, the model performs better after the feature selection method is utilized, increasing by 2.1%, which accords with the heatmap result that indicates the world case has a higher correlation than the South Korea case alone.

Based on the feature importance results shown in Fig. 3, the most significant features for flight, passenger, and cargo are different. DPI and WPUI have a high correlation with airline demand and death cases. In spite of

Table 4.  $R$ -squared value for XG Boost model before and after feature selection

No	Model	Flight	Passenger	Cargo	Average
1	Before	0.931	0.882	0.728	0.847
2	After	0.934	0.920	0.751	0.868
Increment		0.003	0.038	0.023	0.021

this, WPUI\_KOR is relatively less significant.

#### 4.4 Prediction on Domestic and International, Departure and Arrival Demand

The average  $R$ -squared value shown in Table 5, for all sectors (flight, passenger, cargo) shows that domestic demand is less accurate than total demand, indicating that domestic demand is not significantly impacted by the pandemic. However, since domestic demand constitutes only a very small portion of total demand (less than 2%), overall, departure, and

Table 5.  $R$ -squared value for XG Boost model before and after feature selection

No	Model	Flight	Passenger	Cargo	Average
1	Total	0.934	0.920	0.751	0.868
2	Domestic	0.854	0.615	0.647	0.705
3	International	0.932	0.836	0.725	0.831
4	Departure	0.943	0.899	0.587	0.810
5	Arrival	0.942	0.885	0.797	0.875

arrival demand that include domestic demand perform very well. The cargo sector has the poorest accuracy result for departure demand, which indicates that departure demand is less affected by the pandemic.

## V. CONCLUSION

### 5.1 Summary

In spite of this, domestic demand constitutes only a very small portion of total demand (less than 2%), which explains why departure and arrival demand which includes domestic demand performs very well. In the cargo sector, departure demand has the poorest accuracy result, which suggests that the pandemic has had a lesser effect on departure demand.

Air transportation demand is known to be adversely affected by pandemics, specifically the number of flights and passengers; however, cargo is positively affected by them. Pandemic-related features are more negatively correlated with the number of passengers than the number of flights, which indicates that the number of passengers decreased more dramatically than the number of flights.

Additionally, using a feature selection method that excludes only the number of cases and deaths in South Korea (feature names: New\_Cases\_KOR, Cumulative\_Cases\_KOR, New\_Deaths\_KOR, Cumulative\_Deaths\_KOR), performance improvement is shown for the best model, which

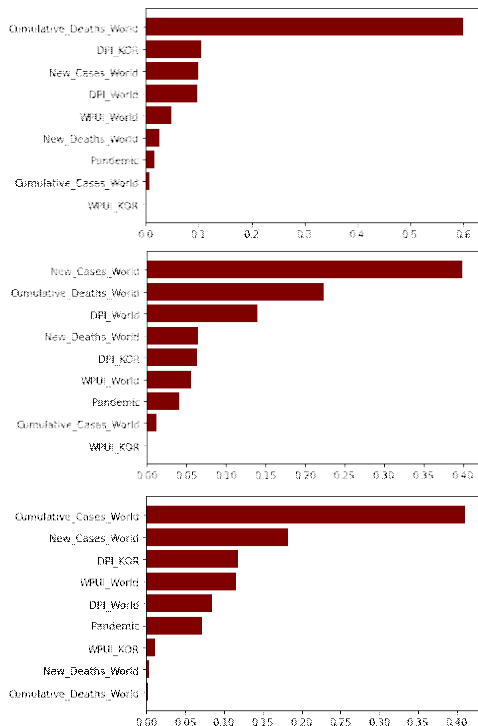


Fig. 3. Feature importance for flight-passenger-cargo demand

indicates that global conditions have more influence on air transportation demand than country conditions. In addition to the number of cases and deaths, WPU and DPI have a strong correlation with air transportation demands. There may be an impact on air travel demand due to uncertainty about the pandemic and discussions about the pandemic.

Analyzing the details of air transport demands by analyzing domestic and international customers, as well as departures and arrivals, reveals that domestic and cargo departure functions less accurately than the total demand, which indicates that demand is less affected by the pandemic.

## 5.2 Future Research

The further research agenda may consider the following aspects for improvement. First, the research only focuses on four major epidemics, including SARS, H1N1, MERS, and COVID-19. As a result, additional information about other pandemics might improve the performance of the model. Second, the pandemic-related features can be enriched with other information, such as global or regional travel policies. However, the information is limited for past pandemics other than COVID-19. Therefore, research focusing on COVID-19 might give a different perspective result. Lastly, the WPU and DPI are built based on Economist Intelligence Unit (EIU) country reports, which might not capture the actual situation for certain countries like South Korea. Therefore, data mining from other resources such as social media might improve the model.

## References

1. Collins-Kreiner N, and Ram Y., "National tourism strategies during the Covid-19 pandemic", *Annals of Tourism Research*, 89, 2021, p.103076, <https://doi.org/10.1016/j.annals.2020.103076>
2. Hu, F., Teichert, T., Deng, S., Liu, Y., and Zhou, G., "Dealing with pandemics: An investigation of the effects of COVID-19 on customers' evaluations of hospitality services", *Tourism Management*, 85, 2021, p.104320, <https://doi.org/10.1016/j.tourman.2021.104320>
3. Bouwer, J., Krishnan, V., and Saxon, S., "Taking Stock of the Pandemic's Impact on Global Aviation", McKinsey & Company, 2022. <https://www.mckinsey.com/industries/travel-logistics-and-infrastructure/our-insights/taking-stock-of-the-pandemics-impact-on-global-aviation> accessed on October 26, 2022.
4. Korea Transport Institute [KOTI], "COVID-19 KOTI Aviation Brief", 2021. [https://www.koti.re.kr/component/file/ND\\_fileDownload.do?q\\_fileSn=301007&q\\_fileId=a533be80-4ab6-4208-9664-7187f2a112f9](https://www.koti.re.kr/component/file/ND_fileDownload.do?q_fileSn=301007&q_fileId=a533be80-4ab6-4208-9664-7187f2a112f9) accessed on October 26, 2022.
5. Park, Y. M., and Jeon, A., "Study on proactive approach against a new large-scale crisis of the aviation and tourism industry such as COVID-19", *Journal of the Korean Society for Aviation and Aeronautics*, 28(4), 2020, pp.176-181. <https://doi.org/10.12985/ksaa.2020.28.4.176>
6. Wang, Y., Wong, C. W. H., Cheung, T. K. Y., and Wu, E. Y., "How influential factors affect aviation networks: A Bayesian network analysis", *Journal of Air Transport. Management*, 91, 2021, p.101995, <https://doi.org/10.1016/j.jairtraman.2020.101995>
7. Kim, Y. R., Lim, J. H., and Choi, Y. C., "Analysis and prospect of export trend of air cargo market before and after COVID-19", *Journal of the Korean Society for Aviation and Aeronautics*, 28(4), 2020, pp.164-170. <https://doi.org/10.12985/ksaa.2020.28.4.164>
8. Kuo, H. I., Chen, C. C., Tseng, W. C., Ju, L. F., and Huang, B. W., "Assessing impacts of SARS and Avian Flu on international tourism demand to Asia", *Tourism Management*, 2002, p.103076

- 29(5), 2008, pp.917-928. <https://doi.org/10.1016/j.tourman.2007.10.006>
9. Loske, D., "The impact of COVID-19 on transport volume and freight capacity dynamics: An empirical analysis in German food retail logistics", *Transportation Research Interdisciplinary Perspectives*, 6, 2020, p.100165, <https://doi.org/10.1016/j.trip.2020.100165>
  10. Kocak, E., Okumus, F., and Altin, M., "Global pandemic uncertainty, pandemic discussion and visitor behaviour: A comparative tourism demand estimation for the US", *Tourism Economics*, 0(0), 2022, pp.1-26. <https://doi.org/10.1177/13548166221100692>
  11. Joo, H., Henry, R. E., Lee, Y. K., Berro, A. D., and Maskery, B. A., "The effects of past SARS experience and proximity on declines in numbers of travelers to the Republic of Korea during the 2015 MERS outbreak: A retrospective study", *Travel Medicine and Infectious Disease*, 30, 2019, pp.54-66. <https://doi.org/10.1016/j.tmaid.2019.05.009>
  12. Ahir, H., Bloom, N., and Furceri, D., "The World Uncertainty Index", SSRN Electronic Journal, 2018. <https://worlduncertaintyindex.com/data/> updated on July 13, 2022, downloaded on September 30, 2022.
  13. Incheon International Airport, "Flight, Passenger, and Cargo Demands Information", Incheon, South Korea. Received on September 29, 2022.