IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 4 March 2024, accepted 27 April 2024, date of publication 1 May 2024, date of current version 10 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3396062

APPLIED RESEARCH

Estimating the Lag Time Between Flight Arrivals and Parking Exit Volumes at a Major Airport: A Practical Approach

ZHIYUAN SHI[1], KE HUANG[2], SHAOZHI HONG[1], AND WANQING SU[3]
[1]Urban Mobility Institute, Tongji University, Shanghai 200092, China
[2]College of Transportation Engineering, Tongji University, Shanghai 201804, China
[3]Institute of Rail Transit, Tongji University, Shanghai 201804, China

Corresponding author: Shaozhi Hong (hongshaozhi@tongji.edu.cn)

This work was supported in part by Shanghai S&T Development Funding under Grant 21DZ1200800, and in part by Shanghai Collaborative Innovation Research Center for Multi-Network and Multi-Modal Rail Transit.

ABSTRACT Predicting the movement of arriving passengers to their landside destination is of great value to airport landside operations. This paper focuses on the arriving passengers who leave the airport by private cars in the terminal parking lots. Disregarding the micro-behavior of passengers, we limit our focus on the time consumption for passengers from flight arrival to parking exit. Traditionally, this information is usually obtained through costly passenger questionnaires. To reduce cost, we develop an alternative way based on time series analysis. Specifically, we try to identify direct causal paths that exhibit significant positive effects, as the lag time to be estimated is the time distribution of these positive lag effects. To overcome the influence of confounding factors, we propose a practical methodology based on developing a set of distributed lag models under different control schemes. The key features of our approach are low data requirements and low mathematical complexity, which make it applicable in the daily operation of airports. We further conduct a case study at Shanghai Pudong International Airport (PVG) to illustrate the proposed methodology. The lag time estimation results are consistent with practical experiences. Sensitivity analyses validate the consistency and reliability of our results. Our research provides a practical way for estimating the lag time between flight arrivals and parking exit volumes, as well as more support for evaluating and improving airport landside operations.

INDEX TERMS Airport landside operations, parking exit volumes, flight arrivals, time series analysis, distributed lag model.

I. INTRODUCTION

The rapid economic expansion and thriving aviation sector in China have led to a significant surge in air travel demand. According to data released by the Civil Aviation Administration of China (CAAC), the entire aviation industry witnessed a remarkable increase in passenger traffic, reaching 660 million passenger traffic in 2019, an increase of over 50% from 436 million in 2015. The continuous growth of air passenger numbers has placed substantial pressure on airports and their infrastructure, demanding a more robust response to meet the escalating requirements [1], [2], [3].

Parking facilities constitute a vital component of airport landside operations, as a considerable number of passengers choose to access and leave airports by private cars such as driving and app-based ride-hailing for convenience. Due to its importance, airport parking has attracted much attention in recent years. Numerous related studies are carried out on fields such as revenue management [4], [5], [6], operations management [7], performance evaluation [8], parking behavior [9], [10], [11], and demand forecast [12], [13], [14], to comprehensively explore and address various aspects of airport parking-related challenges and opportunities.

The associate editor coordinating the review of this manuscript and approving it for publication was Zhe Xiao.

Compared to conventional parking lots, airports exhibit distinct parking demand characteristics influenced by a multitude of factors. Ashford and Wright [15] explained that parking demand at airports is a multifaceted function influenced by several factors, including the volume of individuals accessing the airport, the available modes of access, the characteristics of air travelers, the parking costs, and the duration of parking periods. In the context of airports in China, the separation of drop-off (departure) and pick-up (arrival) channels adds complexity to the parking dynamics. Private cars are required to enter the parking lots to pick up arriving passengers, leading to a surge in parking volumes and causing congestion during peak hours. Estimating the time passengers take from flight arrivals to parking exits is the foundation for congestion relief. This is because the airport operators would be able to predict the peak volumes in the parking lots based on the flight arrival schedules and take measures in advance to alleviate the impact of congestion. The lag time estimates also help evaluate airport service levels, as time indicators are key performance parameters in many previous studies [16]. To achieve the collaborative optimization of airport flight arrival and taxi carrying order, Ding et al. established a matching interaction model based on flight and taxi data from Shanghai Hongqiao Airport [17]. However, to the best of our knowledge, no study has been conducted on flight arrivals and parking exit volumes, which have growing importance due to the increasing usage of online vehicles.

A simple way to estimate the lag time is to divide the distance by the speed. Specifically, divide the average walking distance by the average walking speed and add the average waiting time of each process. This approach is very straightforward, but difficult to obtain information other than the mean value, and its accuracy highly depends on the operators' personnel experience. Another way is considering the behavior of passengers, and developing simulation models to predict their activities, like [18], [19], and [20]. More recently, Wang et al. presented a stochastic model for capturing the decision dynamics of domestic departure passengers, where travel experience and time pressure are considered [21]. Although simulation models can achieve a high level of accuracy and show the distribution of passenger flows, it is costly to tune and adopt them in practice. In addition, simulation models cannot reflect real-time situations. In practice, airport operators often employ passenger questionnaires, such as requesting flight numbers from passengers leaving the parking areas, to assess the time it takes for passengers to travel from flight arrival to parking exit. This time is equivalent to the lag time between flight arrivals and parking exit volumes. The main drawbacks are its high cost and inability to provide real-time information.

To overcome the shortcomings of the above methods, we propose a practical approach based on the distributed lag model to estimate the lag time between flight arrivals and parking exit volumes at PVG. The distributed lag model is a single equation regression technique that can incorporate both lags of the dependent variable and other independent variables to estimate the dynamic causal relationships between variables. In recent years, it has been widely applied in fields such as monetary economics [22], housing prices [23], energies [24], air pollution [25], etc. Moreover, it has gained popularity in transportation research for investigating the relationship between transport and socioeconomic variables (see [26], [27], [28], [29], [30]). Our methodology adopts distributed lag models to depict the dynamic causal relationship between the number of arriving flights and exiting vehicles. The lag time estimates are derived from the results of models under different control schemes. Compared to simulation methods, our methodology does not require a lot of data and arithmetic power. Moreover, it is not mathematically complex, thus very suitable for the daily operations of airports.

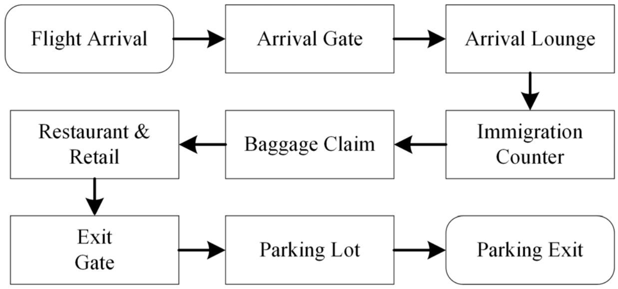


FIGURE 1. Passenger processes from flight arrival to parking exit. Extended based on [16].

The main contributions of this study are as follows. First, we employ causal diagrams to encapsulate our prior knowledge about the underlying impact mechanism and describe the lag time estimation problem as identifying the existence of direct causal paths. Causal diagrams offer simple but powerful tools for modeling causal relationships, which are composed of dots (representing variables) and arrows (representing the direct causal relationship between these variables). The usage of causal diagrams enhances the clarity and reliability of our investigation. Second, we propose a practical approach based on distributed lag models to derive lag time estimates in the presence of multiple known and unknown confounders. A set of distributed lag regressions under different control schemes is introduced to eliminate the impact of confounders since controlling variables cannot block direct causal paths, but indirect and non-causal paths do. Third, the case study at PVG and the following sensitivity analyses not only verify the effectiveness and robustness of our methodology but also provide more insights that may contribute to the improvement of PVG's service quality.

The remainder of this paper is organized as follows. Section II introduces the problem. In Section III, we present the specific framework of our statistical approach. Based on empirical data, Section IV provides the estimation results of the lag time between flight arrivals and parking exit volumes in PVG. Section V performs sensitivity analyses to demonstrate the robustness of the estimation results and the applicability of our methodology. Section VI concludes the study.

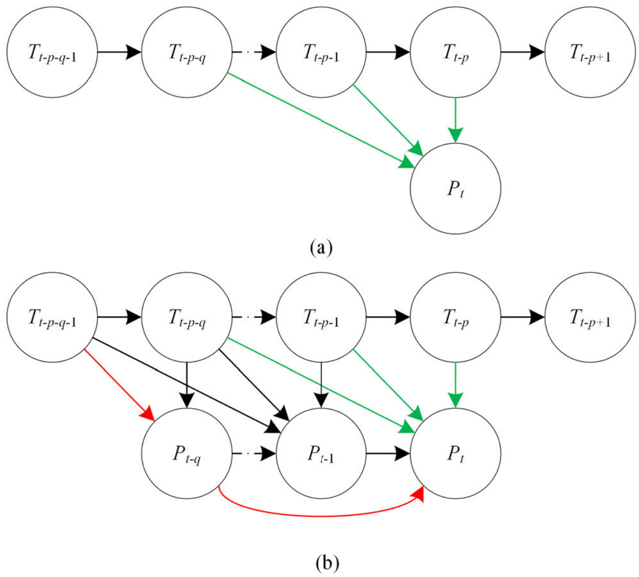


FIGURE 2. Illustration of assumptions and objectives.

II. PROBLEM DESCRIPTION

A. RELATED PASSENGER ACTIVITIES

Fig. 1 depicts the routine activities associated with arrival passengers leaving the airport by private cars. It is important to note that the immigration checking is done only for international passengers.

Flight arrivals are expected to have dynamic causal effects on parking exit volumes, because of the linkage between these processes. That is the effect on parking exit volumes in the future of a change in flight arrivals. The time periods at which this effect acts are the lag time we are trying to estimate.

B. ASSUMPTIONS AND OBJECTIVES

To conduct analysis, the following assumptions are made for the dynamic causal effects based on common sense and practical experiences.

Assumption 1: The effects are positive if exist. This is because an increase in the number of flight arrivals generally results in more passengers heading to the parking lot, leading to a subsequent increase in the number of exiting vehicles.

Assumption 2: The effects will not be immediately following flight arrivals, as passengers require some time to walk to the parking lot for a ride. In other words, it is expected to be a lag effect.

Assumption 3: The effects may last for a period as each passenger takes a different amount of time to complete the relevant process.

Fig. 2(a) illustrates Assumptions 2 and 3 through a simple causal diagram. Let T_t and P_t represent the number of arriving flights and exit vehicles, respectively. According to Assumptions 2 and 3, suppose that the effect lags by p time units and lasts for q, which means that T_{t-p}, T_{t-p-1},..., T_{t-p-q} can directly affect P_t, as indicated by the green arrows in Fig. 2(a).

Our objective is to determine if there is a direct effect from T_{t-m} to P_t for some $m > 0$. For a given m, the existence of such direct effect means that flight arrivals will affect parking exit volumes after $|m|$ periods. These m's multiplied by the length of the time interval are the lag time estimates between flight arrivals and parking exit volumes. It is important to note that when referring to direct effects, we do not imply that T_{t-m} can affect P_t directly without passing through any mediator, which is impossible due to the related passenger activities. In fact, the direct effects we are talking about are relative to the indirect effects that pass through the past values of P_t. As shown in Fig. 2(b), $T_{t-p-q-1}$ can only affect P_t indirectly through the red arrows, so that it is out of the lag time range. Therefore, we need to block these indirect causal paths by controlling the past values of P_t for the accuracy of the lag time estimates.

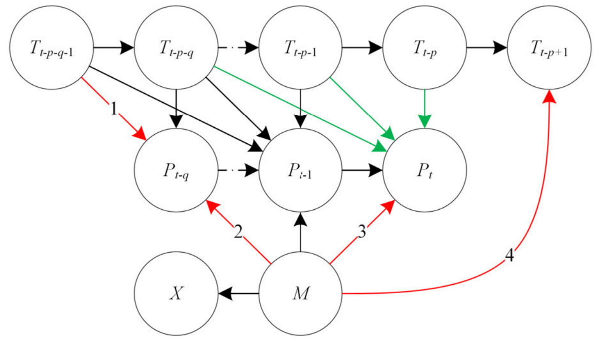


FIGURE 3. Illustration of potential confounding factors.

C. POTENTIAL BIASES DUE TO OTHER FACTORS

It is worth mentioning that controlling the past values of P_t may introduce new challenges when confounding factors like M in Fig. 3 exist.

In the example of Fig. 3, controlling the past values of P_t can effectively block all indirect causal paths but would open a non-causal path that confounds $T_{t-p-q-1}$ and P_t, as shown by the red arrows 1-3. Also note that if M can affect T_{t-p+1}, then T_{t-p+1} would affect $T_{t-p-q-1}$ through a non-causal path via M as red arrows 4 and 3, although there is not a direct causal path from T_{t-p+1} to P_t.

In practice, M might be time series variables like weather conditions and non-time series variables such as airlines. Weather conditions can affect the number of arriving flights and exit vehicles, respectively, as poor weather can lead to flight delays and road congestion around parking lots. The airlines can determine the landing time of arriving flights. Meanwhile, factors such as the level of ground handling services provided by the airlines may affect the volume of parking departures by affecting passengers' waiting and walking time.

Controlling M may not always provide a perfect solution, especially when M is unobservable or unknown. In such cases, one alternative is to find and control a proxy of M, such as X in Fig. 3, which can somewhat mitigate the biases introduced by M. For a multi-terminal airport with multiple parking lots, X can be the number of exiting vehicles of parking lots serving other terminals.

In a word, we need to identify the past values of T_t that have a direct effect on P_t in the presence of multiple known and unknown confounders. Moreover, the positivity of such direct effects is assumed, which is critical in our methodology.

D. PROBLEM DEFINITION

As mentioned earlier, the lag time we aim to estimate corresponds to the temporal distribution of the lag effects from flight arrivals to parking exit volumes. The lag effects represent the direct causal effects from past flight arrivals to the current parking exit volume. The problem of lag time estimation can be transformed into identifying the existence of such causal effects, which are positive, as stated in Assumption 1. Specifically, we aim to determine the presence of direct causal paths from T_{t-m} to P_t for some $m > 0$, rather than estimating the exact value of the causal effects.

Intelligent data-driven methods, such as graph neural networks, have been extensively utilized and have achieved state-of-the-art performance in various traffic forecasting problems [31]. These methods can model complex relationships between time series variables and make accurate predictions. However, their lack of interpretability makes it challenging to directly interpret and understand the relationships between their internal structures and parameters. This limitation can make it difficult to assess whether there are direct causal effects between variables, and thus may not be suitable for our research problem.

In comparison to these "black-box" methods, the traditional linear regression method is much more straightforward and intuitive. It allows for direct interpretation of the impact of each feature on the target variable as the parameters correspond to the weights assigned to each feature. Moreover, linear regression offers advantages in terms of computational efficiency, lower data requirements, and reduced risk of overfitting. Summarizing the aforementioned advantages, we will employ linear regression in our methodology to describe the causal relationship between variables. The presence of direct causal paths is equivalent to the regression coefficient being significantly positive in the linear regression model when all backdoor paths are controlled. Thus, our problem can be transformed into building appropriate linear regression models and identifying lag variables with significantly positive regression coefficients. The lag time corresponding to the eligible lag variables is the estimated lag time between flight arrivals and parking exit volumes.

III. METHODOLOGY

A. GENERAL FRAMEWORK

Fig. 4 shows the flowchart of the proposed method. Raw data, including flight arrival records and parking exit records, are collected from airport information systems. A data cleaning process is performed to eliminate duplicate and exceptional records caused by equipment or other issues. Time series variables representing flight arrivals and parking exit volumes are generated based on the cleaned data and a predefined time unit (e.g., every 10 minutes). We further employ distributed lag regression models to quantify the dynamic effects between flight arrivals and parking exit volumes (see Section III-B).

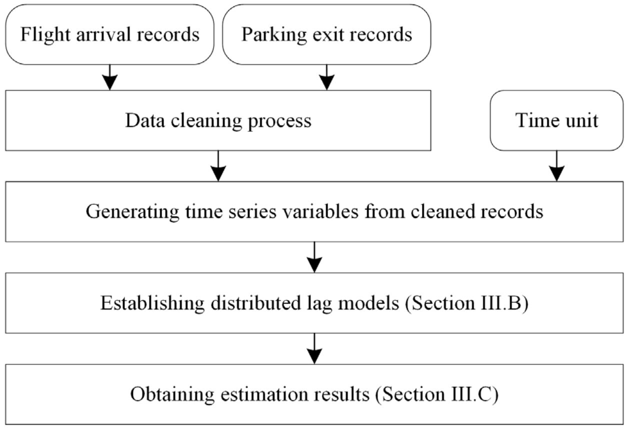


FIGURE 4. Flowchart of the proposed method.

Due to cost and practical constraints, it is impossible to collect all confounders, which may introduce biases in the coefficient estimates of these regressions. To overcome this issue, we will establish models containing different combinations of control variables and compare the results to obtain the lag time estimates. Since indirect and non-causal paths are affected by the choice of control variables, direct ones are not. Based on the positivity of such direct effects, explanatory variables that are positively significant under all settings will be considered directly affecting the dependent variable, and then the lag time estimates will be inferred from these variables (see Section III-C).

B. ESTABLISHING DISTRIBUTED LAG MODELS

The distributed lag model with multiple explanatory and control variables is shown in (1):

$$Y_t = \sum_{i=1}^{l_1} \delta_{1,i} X_{1,t-i} + \ldots + \sum_{i=1}^{l_k} \delta_{k,i} X_{k,t-i} + \sum_{i=1}^{d_1} \beta_{1,i} D_{1,t-i} + \ldots + \sum_{i=1}^{d_r} \beta_{r,i} D_{r,t-i} + \beta_0 \tag{1}$$

where Y_t is the dependent variable; $X_{p,t-i}$'s are the lags of the pth explanatory variable; $D_{q,t-i}$'s are the lags of the qth control variable; k and r denote the number of explanatory and control variables, respectively; δ's and β's are regression coefficients; l's and d's are predetermined lag lengths of explanatory and control variables, respectively.

In our models, the dependent variable is the number of exit vehicles at a parking lot (e.g. P_t), and the explanatory variables are the lags of the number of arriving flights at the terminals it serves (e.g. T_{t-k}). The following decisions are critical before establishing model equations.

- Determine the control variables. The lags of the dependent variable are chosen as control variables to block indirect

TABLE 1. All possible control schemes under two feasible control variables ($r = 2$).

No.	Control $D_{1,t}$?	Control $D_{2,t}$?
1	No.	No.
2	Yes.	No.
3	No.	Yes.
4	Yes.	Yes.

causal paths. If the airport operates more than one parking lot, the lags of the parking exit volumes at other parking lots are also controlled as proxies of confounders.

- Determine the time unit. The quality of our lag time estimates depends on the length of the time unit. On the one hand, rough time units can lead to inaccurate estimation results. On the other hand, exact time units may lead to estimation results sensitive to random errors.
- Determine the lag lengths of explanatory and control variables. This needs to be done in the context of the airport conditions and the time unit. The goal is to ensure that the possible values of the lag time are within the coverage of the model variables.

By taking different combinations of feasible control variables, we will establish several distributed lag models for each pair of parking lots and corresponding terminals. For each pair, the number of equations is 2^r, where r is the number of feasible control variables. For example, if we have two feasible control variables for a pair ($r = 2$), namely $D_{1,t}$ and $D_{2,t}$, four equations will be established corresponding to eight different control schemes as shown in Table 1.

It is worth noting that within our methodology, the dependent variable can also serve as a control variable. This is possible as the lags of the dependent variable can be incorporated on the right-hand side of the distributed lag regression model, commonly referred to as an autoregressive distributed lag (ARDL) model.

C. OBTAINING ESTIMATION RESULTS

We employ ordinary least squares (OLS) to estimate the coefficients of all established regressions. To compute standard errors and determine the statistical significance of variables, we will utilize the heteroskedasticity and autoregression-consistent (HAC) variance estimator proposed by Newey and West [32]. The truncation parameter of the HAC estimator is computed by:

$$m = 0.75N^{1/3} \tag{2}$$

where m represents the truncation parameter, and N denotes the number of observations.

Setting a threshold parameter α (e.g. $\alpha = 0.05$) and selecting the explanatory variables whose significance level reaches α in all regression models. Obviously, their significance needs to be positive. The lag time estimate is obtained by multiplying their lag values and the length of the time unit. For example, if T_{t-3} and T_{t-4} under 10-minute time units are eligible, we can conclude that the lag time estimate is 30-40 minutes.

D. PROOF OF VALIDITY

Here, we demonstrate the validity of the proposed methodology. Specifically, we prove that it can correctly estimate the lag time if Assumption 4 holds.

Assumption 4: Back-door paths have much weaker effects than true causal paths.

A back-door path is any path from X to Y that starts with an arrow pointing to X [33]. The presence of backdoor paths can introduce bias in the regression coefficients. In our problem, paths like the one consists of arrows 4 and 3 in Fig. 3 are typical back-door paths (referred to as type I back-door paths for simplicity). These paths can result in underestimation of the lag time as they may confound the independent variable with the dependent variable before the lag effect begins. While controlling for a confounding variable M can block these backdoor paths, it may create new back-door paths like arrows 1-3 in Fig. 3 (referred to as type II back-door paths). These paths may lead to overestimation of the lag time as they may confound the independent variable with the dependent variable after the lag effect ends.

As mentioned before, in our scenario, M can be time series variables like weather conditions or non-time series variables like airlines. Generally, their impact on parking exit volumes is weaker than the direct effect of flight arrivals. Moreover, for effective airport operations management, meaningful estimates of lag time align with periods when flight arrivals significantly influence parking exit volumes, in which the direct causal effect should prevail. In conclusion, we consider Assumption 4 to be reasonable.

Since M is usually unobservable or unknown, suppose that we are able to control a proxy of M, such as X in Fig. 3. Controlling X can mitigate the impact of M. A "good" choice of X is critical, since we need to effectively mitigate the impact of M to obtain accurate estimates. In this context, we introduce the following Theorem 1.

Theorem 1: If Assumption 1-4 hold, the proposed method can accurately estimate the lag time between flight arrivals and parking exit volumes when controlling X effectively mitigates the impact of M.

Proof: Proving Theorem 1 is equivalent to proving that for each $m > 0$, T_{t-m} is positive significant in all established regressions if there exists a direct causal path from T_{t-m} to P_t. Otherwise, it does not exist. Suppose that the direct causal path exists for $p \leq m \leq p + q$.

i) Consider the case that $p \leq m \leq p + q$:

In this case, T_{t-m} directly affects P_t. T_{t-m} can also indirectly affect P_t through the past values of P_t. Type II back-door paths may exist when controlling M.

Under the control scheme that do not introduce control variables (referred to as scheme I for simplicity), the regression coefficient is positively significant as it captures the combined direct and indirect causal effects.

Under the control scheme that controls the past values of P_t (referred to as scheme II), the regression coefficient remains positively significant as it accounts for the sum of direct effects and the effects of type II back-door paths. According

TABLE 2. Information of datasets.

Datasets	N (raw)	N (cleaned)
Flight Arrival Records at T1	883	585
Flight Arrival Records at TS1	919	792
Flight Arrival Records at T2	1255	936
Flight Arrival Records at TS2	1049	973
Parking Exit Records at P1	53,923	51,148
Parking Exit Records at P2	93,153	89,679

to Assumption 4, the effects of back-door paths cannot negate the positive direct effect.

Under the control scheme that controls X (referred to as scheme III), the regression coefficient remains positively significant as it represents the sum of direct and indirect causal effects, along with a portion of the effects of type II back-door paths.

Under the control scheme that controls both the past values of P_t and X (referred to as scheme IV), the regression coefficient remains positively significant as it captures the combined direct causal effects and a portion of the effects of type II back-door paths.

ii) Consider the case that m$< p$:

In this case, T_{t-m} does not affect P_t directly or indirectly. Type I back-door paths exist when not controlling M. T_{t-m} cannot indirectly affect P_t.

Under scheme III and IV, Type I back-door paths are effectively closed by X (completely if we can control M). Since the regression coefficient captures the effects of type I back-door paths, it cannot be positively significant.

iii) Consider the case that $m > p + q$:

In this case, T_{t-m} does not directly affect P_t. Type II back-door paths exist when controlling M. T_{t-m} can indirectly affect P_t through the past values of P_t.

Under scheme II and IV, Type II back-door paths are effectively closed by X (completely if we can control M). Since the regression coefficient captures the effects of type II back-door paths, it cannot be positively significant. □

IV. CASE STUDY

A. DATA DESCRIPTION

The data used for the analysis are provided by the airport authority, including flight arrival and parking exit records of PVG from Jan. 20, 2023, to Jan. 27, 2023.

PVG consists of four terminals: T1, TS1, T2, and TS2. Among them, TS1 and TS2 are satellite terminals connected to T1 and T2, respectively, via the automated people mover (APM) system. Adjacent to terminals T1 and T2, parking lots P1 and P2 serve as self-park facilities for T1/TS1 and T2/TS2, respectively.

Our raw datasets include the flight arrival records at terminals T1, TS1, T2, and TS2, as well as the parking exit records at parking lots P1 and P2. It is worth mentioning that the raw data exclude the shared flight numbers. Based on our research purpose, we only retain the arrival records of commercial passenger flights and the exit records of non-staff vehicles. Duplicate and exception records due to equipment or other issues are also deleted. Table 2 shows the record counts of each dataset, before and after the data cleaning process.

TABLE 3. Information of the time series variables.

Variables	N	Mean	Std. Dev	Min	Max	Stationary?
$T1_t$	1152	0.51	0.76	0	4	Yes.
$TS1_t$	1152	0.69	0.93	0	5	Yes.
$T2_t$	1152	0.81	1.00	0	5	Yes.
$TS2_t$	1152	0.84	1.03	0	6	Yes.
$P1_t$	1152	44.40	35.75	0	178	Yes.
$P2_t$	1152	77.85	55.77	0	255	Yes.

For each dataset, we generate a time series variable that represents the count of records within 10-minute intervals. The descriptive information of these variables is presented in Table 3. All of them have passed the unit root test proposed by Peter and Perron [34] at a significance level of 0.01, which guarantees their stationarity.

B. MODEL EQUATIONS

Based on PVG's current situation, we construct two sets of distribution lag models with $P1_t$ and $P2_t$ as dependent variables, respectively. The first set takes $T1_t$ and $TS1_t$ as explanatory variables, as shown in (3)-(6), while $T2_t$ and $TS2_t$ are the explanatory variables of the second set, as (7)-(10). Both sets take $P1_t$ and $P2_t$ as feasible control variables. The time unit is taken as 10 minutes and the lag lengths of explanatory and control variables are all set as 10.

$$P1_t = \sum_{i=1}^{10} \delta_{1,i}^{1} T1_{t-i} + \sum_{i=1}^{10} \delta_{2,i}^{1} TS1_{t-i} + \beta_0^1 \tag{3}$$

$$P1_t = \sum_{i=1}^{10} \delta_{1,i}^{2} T1_{t-i} + \sum_{i=1}^{10} \delta_{2,i}^{2} TS1_{t-i} + \sum_{i=1}^{10} \beta_{1,i}^{2} P1_{t-i} + \beta_0^2 \tag{4}$$

$$P1_t = \sum_{i=1}^{10} \delta_{1,i}^{3} T1_{t-i} + \sum_{i=1}^{10} \delta_{2,i}^{3} TS1_{t-i} + \sum_{i=1}^{10} \beta_{1,i}^{3} P2_{t-i} + \beta_0^3 \tag{5}$$

$$P1_t = \sum_{i=1}^{10} \delta_{1,i}^{4} T1_{t-i} + \sum_{i=1}^{10} \delta_{2,i}^{4} TS1_{t-i} + \sum_{i=1}^{10} \beta_{1,i}^{4} P1_{t-i} + \sum_{i=1}^{10} \beta_{2,i}^{4} P2_{t-i} + \beta_0^4 \tag{6}$$

$$P2_t = \sum_{i=1}^{10} \delta_{1,i}^{5} T2_{t-i} + \sum_{i=1}^{10} \delta_{2,i}^{5} TS2_{t-i} + \beta_0^5 \tag{7}$$

$$P2_t = \sum_{i=1}^{10} \delta_{1,i}^{6} T2_{t-i} + \sum_{i=1}^{10} \delta_{2,i}^{6} TS2_{t-i} + \sum_{i=1}^{10} \beta_{1,i}^{6} P2_{t-i} + \beta_0^6 \tag{8}$$

$$P2_t = \sum_{i=1}^{10} \delta_{1,i}^{7} T2_{t-i} + \sum_{i=1}^{10} \delta_{2,i}^{7} TS2_{t-i} + \sum_{i=1}^{10} \beta_{1,i}^{7} P1_{t-i} + \beta_0^7 \tag{9}$$

TABLE 4. Coefficient estimation results of explanatory variables in the models corresponding to parking lot P1. Column (a)-(d) correspond to (3)-(6), respectively. *, **, * represent 1% level, 5% level and 10% level of significance respectively.**

Variables	(a)	(b)	(c)	(d)
$T1_1$	2.19*	0.27	1.29*	0.43
$T1_2$	1.73*	-0.13	0.31	-0.15
$T1_3$	2.19**	0.49	0.46	0.37
$\boldsymbol{T1_4}$	**4.33*****	**2.36*****	**2.39*****	**2.17*****
$\boldsymbol{T1_5}$	**5.43*****	**2.00*****	**2.88*****	**1.81****
$\boldsymbol{T1_6}$	**5.98*****	**1.77*****	**3.68*****	**1.83*****
$T1_7$	5.22***	0.49	3.05***	0.58
$T1_8$	3.31***	-1.31**	0.79	-1.17*
$T1_9$	1.71*	-1.77***	-0.97	-1.77***
$T1_{10}$	0.44	-1.98***	-2.46***	-1.92***
$TS1_1$	1.03	-0.46	-0.16	-0.50
$TS1_2$	0.50	-0.11	-0.38	-0.19
$TS1_3$	1.00	0.44	-0.23	0.27
$TS1_4$	2.06***	1.32**	0.54	1.14**
$\boldsymbol{TS1_5}$	**5.93*****	**4.38*****	**4.30*****	**4.27*****
$\boldsymbol{TS1_6}$	**8.39*****	**4.03*****	**6.44*****	**4.02*****
$TS1_7$	5.59***	-0.37	3.65***	-0.26
$TS1_8$	2.81***	-1.88***	0.71	-1.76***
$TS1_9$	0.77	-2.45***	-1.78***	-2.49***
$TS1_{10}$	0.43	-1.26**	-2.81***	-1.40***

TABLE 5. Coefficient estimation results of explanatory variables in the models corresponding to parking lot P2. Column (a)-(d) correspond to (7)-(10), respectively. $***$, $**$, $*$ represent 1% level, 5% level and 10% level of significance respectively.

Variables	(a)	(b)	(c)	(d)
$T2_1$	1.44	-0.51	-1.22	-0.83
$T2_2$	3.64***	2.69***	1.35*	2.20***
$T2_3$	4.10***	1.51**	1.75**	1.24*
$\boldsymbol{T2_4}$	**5.17*****	**2.49*****	**3.38*****	**2.59*****
$\boldsymbol{T2_5}$	**7.71*****	**3.94*****	**6.07*****	**4.29*****
$\boldsymbol{T2_6}$	**7.41*****	**1.66****	**6.13*****	**2.46*****
$T2_7$	4.98***	-0.88	3.81***	0.30
$T2_8$	4.23***	-1.39*	2.61***	-0.12
$T2_9$	3.41***	-1.83***	1.29*	-0.89
$T2_{10}$	3.40***	-1.35*	0.72	-0.47
$TS2_1$	2.57**	-0.05	-1.10	-0.91
$TS2_2$	2.97***	1.36**	-0.37	0.54
$TS2_3$	1.61*	-0.18	-1.29**	-0.95
$TS2_4$	1.76**	0.85	-0.72	0.11***
$\boldsymbol{TS2_5}$	**5.06*****	**3.61*****	**2.91*****	**3.19*****
$\boldsymbol{TS2_6}$	**8.48*****	**4.53*****	**6.31*****	**4.44*****
$\boldsymbol{TS2_7}$	**7.59*****	**1.73****	**5.66*****	**2.29*****
$TS2_8$	2.67**	-3.52***	0.41	-2.80***
$TS2_9$	1.40	-2.43***	-1.08	-1.81***
$TS2_{10}$	0.62	-3.39***	-2.22***	-2.57***

$$P2_t = \sum_{i=1}^{10} \delta_{1,i}^{8} T2_{t-i} + \sum_{i=1}^{10} \delta_{2,i}^{8} TS2_{t-i} + \sum_{i=1}^{10} \beta_{1,i}^{8} P2_{t-i} + \sum_{i=1}^{10} \beta_{2,i}^{8} P1_{t-i} + \beta_0^8 \tag{10}$$

C. ESTIMATION RESULTS

Tables 4 and 5 present the coefficient estimation results of explanatory variables in the models corresponding to parking lots P1 and P2, respectively. According to (2), the HAC truncation parameter is set to $m = 8$. Taking the threshold parameter as $\alpha = 0.05$, the explanatory variables whose significance level reaches the threshold in all the columns (a)-(d) are highlighted in bold.

TABLE 6. Lag time estimates across different time units.

Terminal	Parking Lot	Units	Significant Lags	Lag time
T1	P1	**5 min**	**9,10,12,13**	**45-65 min**
		10 min	4,5,6	40-60 min
		15 min	3,4	45-60 min
TS1	P1	5 min	9,10,11,12	45-60 min
		10 min	5,6	50-60 min
		15 min	**3,4,5**	**45-75 min**
T2	P2	5 min	7,8,9,10,11	35-55 min
		10 min	4,5,6	40-60 min
		15 min	2,3,4	30-60 min
TS2	P2	5 min	10,11,12,13	50-65 min
		10 min	5,6,7	50-70 min
		15 min	**3,4**	**45-60 min**

As shown in Table 4, $T1_4$, $T1_5$, $T1_6$, $TS1_5$, and $TS1_6$ are positively significant at the 5% significance level under all control schemes. This implies that flight arrivals at terminal T1 have an impact distributed after 40-60 minutes on parking lot P1's parking exit volumes, and flight arrivals at TS1 have an impact distributed after 50-60 minutes on parking lot P1's parking exit volumes. Following the same rule, we can conclude from Table 5 that the lag time estimates are 40-60 minutes and 50-70 minutes, respectively, between terminals T2, TS2, and parking lot P2.

In conclusion, our results indicate that:(1) the lag effect of flight arrivals at T1 on parking exit volumes at P1 distributes between 40-60 minutes; (2) the lag effect of flight arrivals at TS1 on parking exit volumes at P1 distributes between 50-60 minutes; (3) the lag effect of flight arrivals at T2 on parking exit volumes at P2 distributes between 40-60 minutes; (4) the lag effect of flight arrivals at TS2 on parking exit volumes at P2 distributes between 50-70 minutes. We observe that the lag time between satellite terminals and parking exit volumes are longer on average, which may be due to their further distance. Overall, our lag time estimates are consistent with practical experiences.

Moreover, we also note that no explanatory variable is negatively significant under all control schemes. To understand this, consider again that the direct effect is positive, so the negativities must be due to some indirect and non-causal paths. Due to the influence of control variables on these paths, their effects may vary significantly under different control schemes. As a result, the probability that an explanatory variable is negatively significant in all control schemes is low. It also demonstrates the practicality and rationality of our methodology.

V. SENSITIVITY ANALYSIS

We further perform three sets of sensitivity analyses to assess the robustness of the estimation results and get more insights. The first set explores the impact of using different time units. The second set investigates whether the lag time estimates are sensitive to the HAC truncation parameter m. The difference between domestic and international arriving flights is analyzed in the last one.

TABLE 7. Lag time estimates across different truncation parameter settings.

Terminal	Parking Lot	Truncation Parameter	Lag Time
T1	P1	8	40-60 min
		16	40-60 min
		None.	40-60 min
TS1	P1	8	50-60 min
		16	50-60 min
		None.	50-60 min
T2	P2	8	40-60 min
		16	40-60 min
		None.	**30-60 min**
TS2	P2	8	50-70 min
		16	50-70 min
		None.	50-70 min

TABLE 8. Information of additional time series variables.

Variables	N	Mean	Std. Dev	Min	Max	Stationary?
$DT1_t$	1152	0.43	0.70	0	4	Yes.
$IT1_t$	1152	0.08	0.30	0	3	Yes.
$DT2_t$	1152	0.61	0.88	0	4	Yes.
$IT2_t$	1152	0.21	0.46	0	3	Yes.

A. IMPACT OF USING DIFFERENT TIME UNITS

We examine the impact of using different time units, specifically, lag time estimates are derived using 5-minute and 15-minute time units. Note that the lag lengths are set as 20 and 7, respectively. We also tune the HAC truncation parameter according to (2) ($m = 10$ under 5-minute units and $m = 7$ under 15-minute units). Table 5 presents the lag time estimates across different time units.

The estimation results are mostly consistent, except for the matching errors caused by different time units. However, it is still critical to choose proper time units according to the practical situation. As shown in Table 6, the results of TS1 and TS2 under 15-minute units seem inaccurate, while the 5-minute units lead to discontinuous significant lags in T1, which may indicate a potential lack of robustness.

B. IMPACT OF HAC TRUNCATION PARAMETER

We present the lag time estimates across different truncation parameter settings in Table 7. The method used for the estimations is the same as before, except that different HAC truncation parameters are used in calculating the significance level of explanatory variables. Specifically, "none" means we use the general standard errors instead of HAC standard errors.

From Table 7, we can conclude that the lag time estimates are insensitive to changes in the HAC truncation parameter m, as the results are mostly consistent. The only difference is that when we use the general standard errors, the lag time estimates between T2 and P2 change from 40-60 minutes to 30-60 minutes. This points out the reliability of our estimation results and the necessity of introducing HAC estimators.

C. DOMESTIC VERSUS INTERNATIONAL ARRIVALS

It is beneficial to find out if the lag time estimates are sensitive to domestic versus international arrivals, as international arrivals go through a longer process that includes immigration

TABLE 9. Coefficient estimation results of selected explanatory variables. ***, **, * represent 1% level, 5% level and 10% level of significance respectively.

Variables	(a)	(b)	(c)	(d)
$DT1_1$	1.85	-0.22	0.60	-0.09
$DT1_2$	1.21	-0.34	-0.29	-0.34
$DT1_3$	2.04**	0.70	0.57	0.67
$\boldsymbol{DT1_4}$	**4.69*****	**2.89*****	**3.06*****	**2.78*****
$\boldsymbol{DT1_5}$	**5.56*****	**2.13*****	**3.44*****	**2.00*****
$\boldsymbol{DT1_6}$	**5.35*****	**1.24****	**3.44*****	**1.30****
$DT1_7$	4.44***	0.15	2.78***	0.28
$DT1_8$	2.39**	-1.54**	0.38	-1.45**
$DT1_9$	0.33	-2.29***	-1.82**	-2.27***
$DT1_{10}$	-0.64	-1.90***	-3.21***	-1.90***
$IT1_1$	3.69	2.13	4.72**	2.46
$IT1_2$	2.86	0.14	1.95	-0.05
$IT1_3$	2.59	-0.22	-0.28	-0.85
$IT1_4$	2.31	-0.24	-1.66	-0.94
$IT1_5$	4.26**	1.31	-1.21	0.75
$\boldsymbol{IT1_6}$	**9.10*****	**5.30*****	**5.15*****	**5.57*****
$\boldsymbol{IT1_7}$	**10.15*****	**2.95****	**5.98*****	**3.14****
$IT1_8$	8.87***	0.31	5.18***	1.07
$IT1_9$	7.96***	0.12	3.79**	-0.01
$IT1_{10}$	4.20**	-3.73***	0.28	-3.53**

TABLE 10. Coefficient estimation results of selected explanatory variables. ***, **, * represent 1% level, 5% level and 10% level of significance respectively.

Variables	(a)	(b)	(c)	(d)
$DT2_1$	1.33	-0.83	-1.35	-1.17
$DT2_2$	3.41**	2.43***	1.35	2.05***
$DT2_3$	3.98***	1.69**	1.91**	1.45*
$\boldsymbol{DT2_4}$	**5.86*****	**3.37*****	**4.28*****	**3.50*****
$\boldsymbol{DT2_5}$	**8.88*****	**4.76*****	**7.44*****	**5.19*****
$DT2_6$	7.52***	1.12	6.48***	2.10**
$DT2_7$	3.61***	-2.46***	2.63***	-1.15
$DT2_8$	2.42**	-2.75***	1.02	-1.45*
$DT2_9$	2.17**	-2.04***	0.44	-1.16*
$DT2_{10}$	2.90***	-0.80	0.67	0.01
$IT2_1$	0.76	-0.41	-1.76	-0.62
$IT2_2$	3.36	2.93**	0.12	2.07*
$IT2_3$	3.90	1.75	0.50	1.06
$IT2_4$	3.27	0.63	0.25	0.31
$IT2_5$	5.25**	2.49*	2.09	2.29*
$\boldsymbol{IT2_6}$	**8.94*****	**4.75*****	**5.92*****	**4.71*****
$\boldsymbol{IT2_7}$	**11.15*****	**4.87*****	**8.32*****	**5.29*****
$\boldsymbol{IT2_8}$	**11.51*****	**2.99****	**8.13*****	**3.83*****
$IT2_9$	8.61***	-1.23	4.02**	-0.46
$IT2_{10}$	5.78***	-2.52**	0.86	-1.86*

and customs. Based on this consideration, we also classify regional flights as international for research. In our case of PVG, only terminals T1 and T2 had international arrivals during the study period. Time series variables named $DT1_t$, $IT1_t$, $DT2_t$, and $IT2_t$ are generated as shown in Table 8, where "D" and "I" denote domestic and international, respectively.

Two sets of distributed lag models are built following the same settings as (3)-(6) and (7)-(10), respectively. The only difference is that the explanatory variables $T1_t$ in (3)-(6) are replaced by $DT1_t$ and $IT1_t$, so do $T2_t$ in (7)-(10). Tables 9 and 10 present the coefficient estimation results of explanatory variables in the models corresponding to parking lots P1 and P2, respectively.

As shown in Tables 9 and 10, domestic flight arrivals at terminal T1 have an impact distributed after 40-60 minutes on parking lot P1's parking exit volumes; international flight arrivals at T1 have an impact distributed after 60-70 minutes on P1's parking exit volumes; domestic flight arrivals at terminal T2 have an impact distributed after 40-50 minutes on parking lot P2's parking exit volumes; international flight arrivals at T2 have an impact distributed after 60-80 minutes on P2's parking exit volumes. It can be seen that these results are relatively consistent with the previous conclusion of 40-60 minutes and reflect the impact of immigration and customs processes on international arrival passengers.

VI. CONCLUSION AND DISCUSSIONS

In this paper, we conduct a time series analysis based on distributed lag models to estimate the lag time between flight arrivals and parking exit volumes. The main idea of our methodology is transforming the lag time estimation problem into identifying the existence of direct causal effects from the lags of the number of arriving flights to the number of exit vehicles. If the effects exist, the corresponding lag numbers indicate the lag time. In order to eliminate the impact of confounders, a set of distributed lag regression models under different control schemes is introduced. The existence of the effects is recognized when corresponding explanatory variables are positively significant under all control schemes. This approach is practical and convenient as we only need to collect time series corresponding to the number of arriving flights and exit vehicles, which can save a lot of efforts in data collection. In addition, the models and their mathematical calculations are simple and suitable for everyday airport operations.

Taking PVG as an illustrative example, we derive the lag time estimates between terminals T1/TS1 and parking lot P1, as well as T2/TS2 and P2, respectively. The results indicate that the lag time estimates between T1 and P1 are 40-60 minutes; the lag time estimates between TS1 and P1 are 50-60 minutes; the lag time estimates between T2 and P2 are 40-60 minutes; the lag time estimates between TS2 and P2 are 40-60 minutes. Sensitivity analyses are performed to check the robustness of the results, which demonstrate that our results are not only insensitive to different choices of the HAC truncation parameter m but also consistent under different time units. 10 min is suggested to use as time units for a balanced trade-off between accuracy and robustness. Moreover, we also find that our methodology can not only provide lag time estimates that are consistent with practical experiences but also identify the distinction between international and domestic arrivals due to immigration and customs.

In a word, our research offers a practical methodology for estimating the lag time between flight arrivals and parking exit volumes. The accurate lag time estimates are helpful not only in automating landside operations like staffing parking lots and TNC dispatch based upon real-time flight arrivals data but also in evaluating the efficiency and service quality of the airports. Some directions for future research are summarized as follows. First, it is beneficial to adopt our methodology on more transportation modes, such as taxis and public transit, to gain more insights into airport landside operations. Second, it is worthwhile to use pedestrian simulation technology to verify our findings, as the lag time is highly related to the movement of passengers. Third, gathering datasets from other airports and time periods can help further validate our methodology and make necessary improvements.

ACKNOWLEDGMENT

The authors sincerely thank the associate editor and two anonymous reviewers for their valuable suggestions and comments, which have greatly contributed to the improvement of this manuscript.

REFERENCES

[1] A. Evans and A. Schäfer, "The impact of airport capacity constraints on future growth in the U.S. air transportation system," *J. Air Transp. Manage.*, vol. 17, no. 5, pp. 288–295, 2011.

[2] L. Budd, S. Ison, and T. Budd, "Improving the environmental performance of airport surface access in the U.K.: The role of public transport," *Res. Transp. Econ.*, vol. 59, pp. 185–195, Nov. 2016, doi: 10.1016/j.retrec.2016.04.013.

[3] C. Malandri, L. Mantecchini, and M. N. Postorino, "Airport ground access reliability and resilience of transit networks: A case study," *Transp. Res. Proc.*, vol. 27, pp. 1129–1136, Jan. 2017, doi: 10.1016/j.trpro.2017.12.022.

[4] A. Kanafani and L. H. Lan, "Development of pricing strategies for airport parking—A case study at San Francisco Airport," *Int. J. Transp. Econ.*, vol. 15, no. 1, pp. 55–76, 1988. [Online]. Available: http://www.jstor.org/stable/42748215

[5] A. Papayiannis, P. V. Johnson, D. Yumashev, and P. Duck, "Revenue management of airport car parks in continuous time," *IMA J. Manage. Math.*, vol. 30, no. 1, pp. 1–35, Jan. 2019, doi: 10.1093/imaman/dpy015.

[6] C. Cheng and P. Qi, "Impact analysis of parking price adjustment on the quality of service of airport parking lots for light vehicles," *J. Adv. Transp.*, vol. 2019, pp. 1–9, Jun. 2019, doi: 10.1155/2019/3847837.

[7] *Handbook to Assess the Impacts of Constrained Parking at Airports*, Nat. Academies Sci. Med., 2010.

[8] X. Liu, Q. Zhang, and Y. Lu, "Evaluation system of hub airport parking building project based on fuzzy comprehensive evaluation method," in *Proc. 7th Int. Conf. Inf. Sci., Comput. Technol. Transp.*, May 2022, pp. 1–11.

[9] Y. Ge, A. Biehl, S. Ravulaparthy, V. Garikapati, M. Lunacek, and C. Phillips, "Joint modeling of access mode and parking choice of air travelers using revealed preference data," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2675, no. 11, pp. 699–713, Nov. 2021, doi: 10.1177/03611981211019037.

[10] M. K. Jha, P. Schonfeld, and F. McCullough, "A machine learning and simulation-based dynamic parking choice model for airports," *J. Air Transp. Manage.*, vol. 111, Aug. 2023, Art. no. 102425, doi: 10.1016/j.jairtraman.2023.102425.

[11] H. Qin, J. Gao, G. Zhang, Y. Chen, and S. Wu, "Nested logit model formation to analyze airport parking behavior based on stated preference survey studies," *J. Air Transp. Manage.*, vol. 58, pp. 164–175, Jan. 2017, doi: 10.1016/j.jairtraman.2016.10.011.

[12] R. Guo, X. Shen, and H. Kang, "Improved CS algorithm and its application in parking space prediction," *J. Bionic Eng.*, vol. 17, no. 5, pp. 1075–1083, Sep. 2020, doi: 10.1007/s42235-020-0056-x.

[13] C.-W. Lin, W. Li, Y. Huanh, and G. Yang, "Parking demand forecasting in airport ground transportation system: Case study in Hongqiao Airport," *Inf. Sci.*, 2015.

[14] F. Adam and M. Vanderschuren, "Determining long-term parking needs at OR Tambo international airport in Johannesburg," *J. Air Transp. Manage.*, vol. 15, no. 6, pp. 363–367, Nov. 2009, doi: 10.1016/j.jairtraman.2009.04.007.

[15] N. Ashford and P. H. Wright, "Airport access," in *Airport Engineering*, 1992, pp. 418–444.

[16] A. Thampan, K. Sinha, B. R. Gurjar, and E. Rajasekar, "Functional efficiency in airport terminals: A review on overall and stratified service quality," *J. Air Transp. Manage.*, vol. 87, Aug. 2020, Art. no. 101837, doi: 10.1016/j.jairtraman.2020.101837.

[17] X. Ding, Z. Liu, G. Shi, H. Hu, J. Chen, K. Yang, S. Wan, and J. Wu, "The optimization of airport management based on collaborative optimization of flights and taxis," *Discrete Dyn. Nature Soc.*, vol. 2022, pp. 1–16, Feb. 2022, doi: 10.1155/2022/4542299.

[18] S. Eilon and S. Mathewson, "A simulation study for the design of an air terminal building," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, no. 4, pp. 308–317, Jul. 1973, doi: 10.1109/TSMC.1973.4309241.

[19] M. R. Gatersleben and S. W. Van der Weij, "Analysis and simulation of passenger flows in an airport terminal," in *Proc. Winter Simulation Conf. Simulation, Bridge Future*, 1999, pp. 1226–1231, doi: 10.1109/wsc.1999.816845.

[20] P. Fonseca i Casas, J. Casanovas, and X. Ferran, "Passenger flow simulation in a hub airport: An application to the Barcelona international airport," *Simul. Model. Pract. Theory*, vol. 44, pp. 78–94, May 2014, doi: 10.1016/j.simpat.2014.03.008.

[21] Y. Wang, T. Liu, M. Hu, S. Alam, and V. N. Duong, "Temporal patterns underlying domestic departure passengers behavior in the airport," *IEEE Access*, vol. 8, pp. 127969–127980, 2020, doi: 10.1109/ACCESS.2020.3008438.

[22] J. U. Matola, "How monetary policy affects industrial activity in Malawi: Evidence from ARDL and VAR models," *Cogent Econ. Finance*, vol. 11, no. 1, Dec. 2023, doi: 10.1080/23322039.2023.2190643.

[23] F. Chege, H. F. Gholipour, and S. Yam, "The long-run impact of remittances on house prices in Kenya," *Int. J. Housing Markets Anal.*, vol. 2023, pp. 1–15, Jun. 2023, doi: 10.1108/ijhma-04-2023-0047.

[24] M. D. Doran, M. M. Poenaru, A. L. Zaharia, S. Vătavu, and O. R. Lobonţ, "Fiscal policy, growth, financial development and renewable energy in Romania: An autoregressive distributed lag model with evidence for growth hypothesis," *Energies*, vol. 16, no. 1, p. 70, Dec. 2022, doi: 10.3390/en16010070.

[25] A. Abedi, M. M. Baygi, P. Poursafa, M. Mehrara, M. M. Amin, F. Hemami, and M. Zarean, "Air pollution and hospitalization: An autoregressive distributed lag (ARDL) approach," *Environ. Sci. Pollut. Res.*, vol. 27, no. 24, pp. 30673–30680, Aug. 2020, doi: 10.1007/s11356-020-09152-x.

[26] D. He, Q. Yin, M. Zheng, and P. Gao, "Transport and regional economic integration: Evidence from the Chang-Zhu-Tan region in China," *Transp. Policy*, vol. 79, pp. 193–203, Jul. 2019, doi: 10.1016/j.tranpol.2019.04.015.

[27] M. Z. Khan, "Analyzing road transport (Passenger and Freight) demand in Pakistan with auto-regressive distributed lag co-integration approach," *Transp. Res. Record: J. Transp. Res. Board*, vol. 2675, no. 4, pp. 155–170, Apr. 2021, doi: 10.1177/0361198120974007.

[28] S. Bektaş, "Examining the impact on road safety performance of socioeconomic variables in Turkey," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2676, no. 1, pp. 435–445, Jan. 2022, doi: 10.1177/03611981211036374.

[29] M. Shafique, A. Azam, M. Rafiq, and X. Luo, "Investigating the Nexus among transport, economic growth and environmental degradation: Evidence from panel ARDL approach," *Transp. Policy*, vol. 109, pp. 61–71, Aug. 2021, doi: 10.1016/j.tranpol.2021.04.014.

[30] J. Chi, "Asymmetric gasoline price effects on public transit ridership: Evidence from U.S. cities," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2676, no. 5, pp. 643–659, May 2022, doi: 10.1177/03611981211069065.

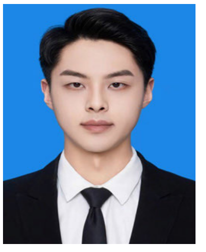
[31] W. Jiang and J. Luo, "Graph neural network for traffic forecasting: A survey," *Expert Syst. Appl.*, vol. 207, Nov. 2022, Art. no. 117921, doi: 10.1016/j.eswa.2022.117921.

[32] W. K. Newey and K. D. West, "Hypothesis testing with efficient method of moments estimation," *Int. Econ. Rev.*, vol. 28, no. 3, p. 777, Oct. 1987, doi: 10.2307/2526578.

[33] J. Pearl and D. Mackenzie, *The Book of Why: The New Science of Cause and Effect*. Basic Books, Inc., 2018.

[34] P. C. B. Phillips and P. Perron, "Testing for a unit root in time series regression," *Biometrika*, vol. 75, no. 2, p. 335, Jun. 1988, doi: 10.2307/2336182.

ZHIYUAN SHI received the B.S. degree in traffic engineering from Tongji University, in 2022, where he is currently pursuing the Ph.D. degree in urban mobility. His research interests include air transportation, transportation operations management, and urban logistics.



KE HUANG received the B.S. degree in transportation engineering from Fuzhou University, in 2021. He is currently pursuing the M.S. degree in transportation planning and management with Tongji University. His research interests include data analysis and optimization of parking management.

SHAOZHI HONG received the B.S. degree in materials science and engineering, the M.S. degree in architectural and civil engineering, and the Ph.D. degree in industrial engineering and management from Tongji University, in 1995, 2002, and 2011, respectively. He is currently the Director's Assistant with the National Maglev Transportation Engineering Research and Development Center and the Deputy Director of Shanghai Collaborative Innovation Research Center for Multi-Network and Multi-Modal Rail Transit. His research interests include airport economy, mobility planning, infrastructure digitization, and security emergency management. He is a member of Shanghai Municipal Transportation Commission and Shanghai Urban and Rural Construction Committee.

WANQING SU received the B.S. degree in transportation from Southwest Jiaotong University, in 2022. She is currently pursuing the master's degree in transportation with the Institute of Rail Transit, Tongji University. Her research interests include profit sharing and incentive mechanisms in freight air-rail intermodal transport.