Hindawi

Research Article

Wireless Network-Aided Delay Information System Correlation with Airport Grid Distribution Based on Multideterminants Big Data

Yujie Yuan and Xiushan Jiang 🆔

School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

Correspondence should be addressed to Xiushan Jiang; xshjiang@bjtu.edu.cn

The intelligent sensing and communication technology in the airports' grid information system provides a multidimensional big data set for analyzing flight delays. These data from air traffic control, weather, and multiple determinants will cause initial flight delays. Due to the influence of adjacent flight time correlation, the initial delay causes the delay of subsequent flights, discovered by mining information sensing data, forming the phenomenon of flight delay diffusion. Different determinants will lead to the delay diffusion form of different regions, and more seriously, it will lead to "disaster area" delay in the whole regional grid information structure. To analyze the spatial impact of each factor on flight delay and explore the regional distribution of delay determinants, this paper combined the spatial regression model and determined the key explanatory variables by statistical and processing of the aviation system data. The case study showed the spatial airport delay characteristics in terms of aircraft movements in China. After processing intelligent sensing and communication data, the results show that there is a spatial effect between airports in terms of delay and determinants. The high-delay clusters of delay constraints principally occurred in the Beijing-Tianjin-Hebe and Yangtze River Delta urban agglomerations. Direct flights, weather, new flight routes, take-off, and landing capacity have a more critical impact on spatial airport delays. The use of Internet of Things technology to perceive, analyze, and integrate multiple information of airport delay and combine spatial analysis models can accurately mine delay characteristics and effectively achieve digital and intelligent flight delay management.

1. Introduction

Intelligent flight sensing and communication data can extract multidimensional data about aircraft flight trajectories, weather, airport operations, and air capacity. Based on these data, systematic analysis of the delay information in the airport grid is the key to reducing the direct loss of passengers and carriers and promoting the economic development of the civil aviation region. Delayed spatial grid analysis combines the idea of the Internet of Things (IoT) network technology with the current research trend. This is achieved by integrating passenger mobile data, aircraft operation information, meteorological distribution information, and airport status information. Based on the analysis of flight information, Chen et al. calculated the indirect economic impact of flight delays on the Chinese economy and concluded that the total indirect influence in 2013 was 355.71 billion RMB, which also stressed the importance of controlling flight delays [1]. Except for China, the Federal Aviation Administration (FAA) has proposed that the increase in flight delays endowed tremendous pressure on the US air travel system with billions of dollars loss of airlines, passengers, and society annually. In 2007, the economic losses caused by airlines amounted to 8.3 billion dollars, and the losses included increased staff, fuel, and maintenance costs [2, 3]. Also, Air Traffic Flow Management (ATFM) estimated the total cost of delays in Europe (including all causes and reaction costs) to be 1.15 billion euros in 2011. The average delay cost for delayed flights has already reached 1,660 euros [4]. Flight delays have brought a significant impact on the global economy, emphasizing the importance of controlling flight delays.

Except for the enormous economic cost loss, the direct consequence of flight delays has reduced the on-time performance, which is also a widespread concern in the civil aviation industry. In 2017, 2.895 million China's passenger airlines were on time (4.039 million flights in total), and the average flight rate was 71.67%. Several primary reasons for the flight delay include air carrier problems, extreme weather, and air traffic control [5, 6]. Beyond that, spatial correlation can also lead to the propagation effects of airport delays. By using machine learning to simulate the aircraft operation, delays in the flight of one aircraft can affect the subsequent flight, which ultimately leads to delays in the propagation of airport pairs [7, 8]. By integrating the sensor communication information among the airports, routes, and aircraft, the airport delays have formed an extensive delay grid information system. How to accurately analyze the spatial relationship between multiple airports in the delay grid and lucubrate their determinants has become significant and challenging issues for civil aviation delay.

At present, the combination of new technologies such as the Internet of Things, big data, artificial intelligence, and 5G communication with airport operation management has begun to receive attention [9–11]. Among them, the intelligent analysis of delay management and control is also one of the research contents using the new scientific and technological revolution and industrial reform. Analyzing and integrating various information and element resources of a flight delay can help finally realize digital and intelligent decision-making on delay prevention, control, and mitigation. It is the extensive application and deep integration of new-generation technologies such as the Internet of Things, big data, and artificial intelligence in delay analysis. In addition to intelligent data mining, an effective analysis model is also the key to studying the distribution characteristics of delay in the airport grid.

About correlation between delay information system and spatial airport grid distribution, Hansen and Hsiao used the econometric model to examine the daily mean of 32 airports' take-off delays in the United States from a time dimension. The trend effects, including aircraft queues, flight schedule, and meteorological conditions, are statistically analyzed [12]. They found that the increase in total flight and operation demand would aggravate airport delays in the airport grid distribution. The delay effect of the destination and route weather counted on the number of flights. Zou et al. conducted a comprehensive empirical analysis on the impact between flight delays and flight frequencies in the US air transport system [13]. The results showed that flight frequency had a positive impact on flight delays. Duran-Fernandez and Santos found four critical variables that can explain delays in European airports (market concentration, coordination, hub airports, and hub airlines) [14]. In Europe, although the flight delay at the hub airport was higher than that of the nonhub airport, the flight delay of the hub airline was lower than that of the nonhub airline, which explained why the spoke-type hub system in Europe was not comprehensive, and the degree of control about the take-off and landing of airports was incomplete. Lall first attempted to use the count regression model to investigate delays and

delay determinants among the three airports in New York City [15]. The Poisson regression model and the least-squares regression model were used to analyze the influencing factors of New York airport delay, while severe weather had the most significant impact on expected delays.

Since the relevant studies are interdisciplinary, scholars have used various parameters or nonparametric methods in their research. However, the study on the airport delays from the spatial grid is scarce, fragmentary, and unmethodical. Therefore, the contributions of this paper include (1) exploring the spatial grid pattern of flight delay at the city level; (2) evaluating the comprehensive spatial autocorrelation of delays between airport grids; and (3) quantitatively identifying the geographical distribution characteristics of each delay determinant and calculating its impact degree by processing the flight sensing and communication data. This paper uses spatial regression models to analyze the correlation and determinants of delays among multiple airports, and the results can provide a reference for the focus of delay prevention and control in different regions in air traffic management.

The results and multivariables statistics methods are reported in Section 2, which also show the relevant explanatory variables in this analysis. Section 3 presents the method of modeling. Section 4 discusses the methodology, and Section 5 provides conclusions and policy recommendations.

2. Multivariables Determination by Big Data Mining

The Internet of Things and big data application needs to be implemented based on the various flight delay activities and operations. Figure 1 shows the elements of the airport grid operation and correlation. According to these elements, we propose an airport grid of delay element framework based on the Internet of Things technology, which divides into three categories (operation control, aviation meteorology, and collaborative interaction). In Figure 2, the three categories include flight plans, aircraft track, aircraft performance, operation rules, flow control data, meteorological data, airport collaboration, company collaboration, passenger collaboration, and other determinants.

The intelligent sensing and communication data of airport delay IoT structure processed were from the Civil Aviation Administration of China, the Statistics Bureau, and the Beijing Capital International Airport database with route conditions and corresponding weather information system data. However, the weather database is an hourly record, and the flight database is not always consistent with other items. Therefore, in data processing, the database is divided into route (including real time flight route monitoring, spatial positioning and tracking, passenger movement, and other data) and weather type. According to the 2017 Civil Aviation Development Statistics Bulletin issued by the Civil Aviation Administration of China, there are ten major airlines i.e., Guide Air, Air China, China Eastern Airlines, Hainan, Shenzhen, Sichuan, Xiamen, Shandong, Shanghai, and Tianjin Airlines. The executed
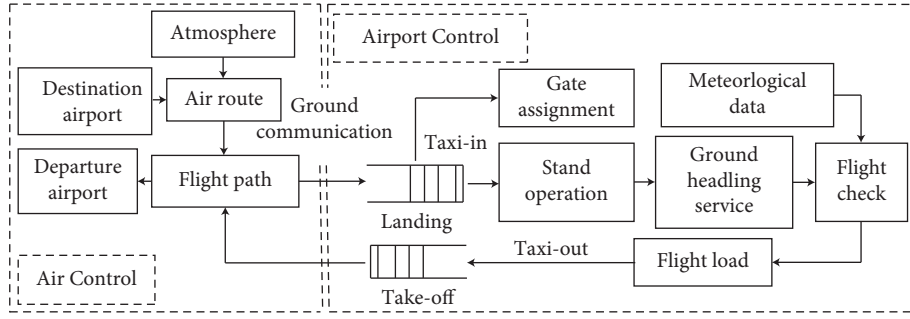
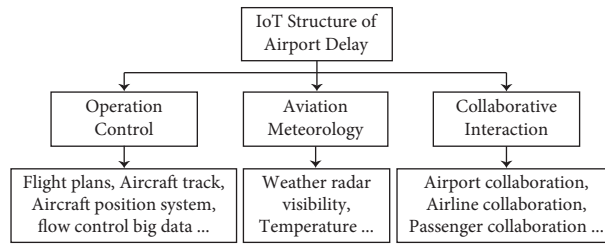Figure 1: The elements of the airport grid operation and correlation.

Figure 2: The IoT structure of airport delay.

flight volume accounted for 78% of the total flight volume, but its average flight on-time performance was 66.9%, which did not reach the average standard rate of all flights in China. The top 100 airports have almost covered the whole major airlines and routes under the comprehensive Chinese delay system. Therefore, the data are driven from top 100 traffic volume and processed airports from January 1, 2017, to December 31, 2017, including airport flight conditions, aircraft type, delay time, passenger load factor, and weather at the corresponding time. The data of the collection are as follows.

2.1. Dependent Variable from IoT Structure of Airport Delay. In the delay information system, the dependent variable is the average delay time for 15 minutes or more relative to the scheduled arrival/departure times. The Civil Aviation Administration of China has demonstrated the delays of airlines by the delaying of flights. In 2016, according to the draft "Statistical Measures for the Regular Flight of Civil Aviation (Consultation Draft)," the on-time flight referred to the flight [16] that arrived at least 15 minutes (inclusive 15 minutes) after the scheduled arrival time. The standard turn time was set based on the airport passenger throughput published by the Civil Aviation Authority in the previous years, stipulating the maximum time from the withdrawal of the airport to take-off. Table 1 contains the standard airport turn time. Therefore, during the sample period of the study, the calculation formula (1) and formula (2) for the delay time of each flight in each airport are as follows:

$$AD_{ij} = ATA_{ij} - ETA_{ij} - Taxiing_{ij}, \quad (1)$$

$$DD_{ij} = ATD_{ij} - ETD_{ij} - Taxiing_{ij}. \quad (2)$$

The total average time is as follows:

$$TD_i = \sum_{i=1, j=1}^{n} AD_{ij} + \sum_{i=1, j=1}^{m} \frac{DD_{ij}}{m+n}, \quad (3)$$

where AD_{ij}, $AD_{ij} > 0$ indicates the arrival delay time of flight j in airport i, ATA_{ij} means the actual arrival time, ETA_{ij} represents the estimated arrival time, DD_{ij}, $DD_{ij} > 0$ symbolizes the departure delay time of flight j in airport i, ATD_{ij} intimates the actual departure time, ETD_{ij} is the estimated departure time, and $Taxiing_{ij}$ indicates the standard turn time. TD_i implies the average delay time of airport i, n is the total number of flights arriving, and m signifies the total number of flights departing.

2.2. Independent Variable from IoT Structure of Airport Delay. The concept of the IoT is to connect any object with the network. Objects exchange and communicate information through information dissemination media to achieve intelligent identification, positioning, tracking, supervision, and other functions. Bringing the idea of the IoT into the analysis of flight delay can help research on mining information data related to delay from the systems perspective and as independent variables. Independent variables include average passenger throughput for each shift and average take-off and landing times for all aircraft at the airport each day, which are from the flight database of the Civil Aviation Administration, intelligent airport sensing, communication datasets, and summarized by shifts and hours.

Due to the inconsistent minimum spacing required between aircraft pairs, aircraft take-off and landing can affect airport capacity, resulting in flight delays, especially when instrument conditions are in effect [17–19] through

TABLE 1: Standard airport turn time.

Airport size	Standard ground turn time (min)
Domestic airport and overseas airport with annual passenger throughput ≥ 20 million passengers	30
Domestic airport with annual passenger throughput ≥ 10 million passengers	25
Domestic airport with annual passenger throughput ≥ 5 million passengers	20
Domestic airport with annual passenger throughput < 5 million passengers	15

aircraft positioning data. Aircraft types will also affect airport delays, especially in heavy-duty operations where mixed take-off (landing) has the most significant adverse impact.

Considering the capacity and airspace constraints in terms of airspace, the number of flights can reflect the congestion of the airspace. Duestablished delay causality grid (DCG) based on the Granger causality test and determined the airports associated with the delayed propagation links of the airports [20]. Figure 3 shows the directed grid which includes building DCGs and counting the number of flights between the two airports to reflect the number of airspace congestion routes at each airport.

In terms of airlines, through flight and passenger tracking data, considering the actual average capacity, the number of direct flights from each airport and the number of new direct flights to and from the port are independent variables. The increased demand (obtained through the movement trajectory of passengers' mobile information) has imposed the busy degree of airports and airlines, especially the hub airports. The average airport capacity will directly affect the air traffic congestion, thus affecting the take-off and landing time of aircraft, so it is necessary to consider the number of direct flights from each airport since direct flights are the decisive factor that can directly lead to the delay of the next flight. At the same time, the continued high increasing for new air routes has complicated the crowded airspace structure.

The weather system is an indispensable factor in the "IoT network" of airport delays. Lousy weather conditions can lead airport dysfunction and cause delays in almost all operation phases. Besides, due to adverse weather events, airport visibility will be reduced, resulting in large-scale airport delays. Previous studies analyzed the visibility in detail and collected various weather elements that affect airport visibility [20–22]. In the delay information system, regarding the daily weather conditions of each airport meteorological bureau (such as the visibility affecting aircraft take-off and landing), he sorted out the weather conditions of re-air flight and ground take-off and determined the factors that would affect the delay. All subweather variables would integrate into a total weather variable. The counting method is as shown in formula as follows:

$$WEA = \left\{ \begin{array}{l} 1, \text{if delay occurs by weather,} \\ 0, \text{otherwise} \end{array} \right\}. \quad (4)$$

According to the meteorological radar big data from airport grid IoT structure, the conditions for weather selection are as follows [13, 25]:

(i) If there is a severe thunderstorm reported within 50 miles of the airport, the indicator variable will take a value of 1; otherwise, it is zero.

(ii) In the route, there are moderate and heavy road thunderstorms with a value of 1, otherwise zero.

(iii) The airport has heavy snow (24 hour snowfall between 5.0 and 10 mm) or blizzard (24 hour snowfall above 10 mm) with a value of 1, otherwise zero.

(iv) There is heavy rain at the airport (precipitation with a rainfall of more than 16 mm per hour, or a continuous rainfall of more than 30 mm for 12 hours, or precipitation with a rainfall of more than 50 mm for 24 hours). The value is 1; otherwise, it is zero.

(v) Strong winds will appear at the airport (is above level 4) with a value of 1, otherwise zero.

(vi) Haze weather at the airport (greater than 80%) has a value of 1, otherwise zero.

(vii) When the airport cloud level is lower than the lowest decision height (10 meters) of the instrument landing level, the value is 1; otherwise, it is zero.

(viii) The sandstorm storm at the airport (less than 1 km) takes the value 1, otherwise zero.

With the above research on the relationship between airport delays and related factors, the relevant variables of nine airport delays have been obtained based on intelligent sensing and communication data collation and statistics (as shown in Table 2).

3. Methodology

Spatial effects and autocorrelation tests must be carried out firstly on critical variable data before modeling in the delay information system. If a spatial effect exists, a spatial regression model will be further constructed to reach estimated measurement.

3.1. Spatial Correlation between Airport IoT and Other Independent Variables. Before studying the spatial correlation between airport IoT and other independent variables, it is necessary to determine the spatial correlation of delay between the airport grid. In order to detect the spatial relationship between delays, it is necessary to carry out a Moran I index test among multiple airport pairs for the average concentration of airport delays in spatial units and test the similarity, difference, or independence of airport delays across China.
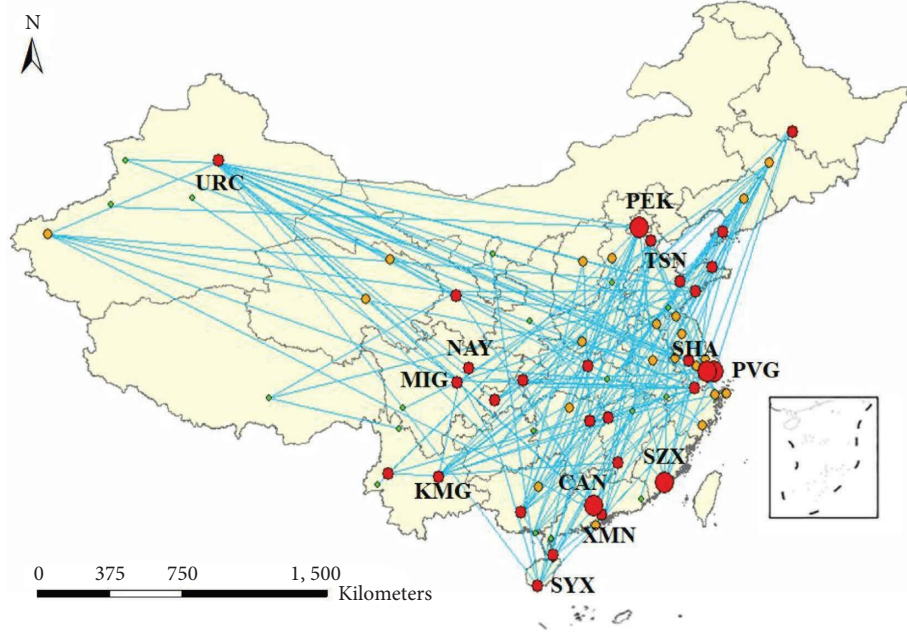
Figure 3: Delayed causal grid of the Granger causality test.

Table 2: A summary of key variable statistics of airport delay IoT structure.

Variable	Description	Min	Max	Mean	Std. dev.
TD	Airport average delay time (min)	18.31	49.22	32.57	7.01
AAC	Actual airport average capacity (seat/flight)	40.40	86.80	66.90	9.34
DFC	Number of direct flights from each airport (count)	16.86	1636.33	253.49	310.98
NDF	Number of new added flights to and from each airport (count)	114.00	1499.00	163.73	135.73
PTD	Average number of passengers arriving (departing) at each airport (person/flight)	4.00	247.00	57.19	50.38
ATL	The average number of landings of the total aircraft arriving (departing) at the airport every day (count)	0.00	32.00	10.78	7.88
NHA	Airport heavy aircraft ownership (count)	149.00	333.00	238.66	39.08
ACR	Airspace congestion route (count)	0.00	1619.00	144.80	248.62
WEA	Delay weather conditions (count)	1.00	29.00	5.89	6.15

Moran's I is ranging from −1 to 1, where if the value > 0, it indicates a positive correlation, as a high (or low) delay airport is adjacent to a high (or low) delay airport. Value < 0 indicates a spatial negative variable correlation, indicating that a high-delay airport is adjacent to a low-delay airport; when the index is equal to 0, there is no spatial relationship between airports. The higher the spatial correlation between airports, the greater the absolute value of the index will be. Equations (5)–(7) is as follows:

$$I = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (x_i - \overline{x})(x_j - \overline{x})}{S^2 \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}}, \tag{5}$$

$$S^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i - \overline{y}), \tag{6}$$

$$\overline{y} = \frac{1}{n} \sum_{i=1}^{n} y_i, \tag{7}$$

where I represents Moran's I, n is the number of airports in a Geospatial space, x_i and x_j are the delay values of airport i and airport j, respectively, w_{ij} is an element of the space weight matrix, and $\overline{x}$ is the average of all observations for an attribute feature, x, in n study areas.

On the other hand, in the airport grid, the phenomenon of delay accumulation between local airports is measured by the Moran scatter plots and LISA aggregate plots, which can reflect the degree of association between an airport and its neighboring airports. The Moran scatter plot is calculated as follows:

$$I_I = \frac{(x_i - \overline{x}) \sum_{j=1}^{n} w_{ij}(x_j - \overline{x})}{S^2}. \tag{8}$$

The LISA aggregation graph enables visualization operations to be performed directly on the map by delaying aggregation. It has pronounced and intuitively showed the spatial distribution of delays situation in the located area.

3.2. Spatial Weight Matrix by Route Tracking Position Data. According to route track positioning data of aircraft, different from spatial neighboring weight matrices, airports cannot directly determine whether they are contiguous. The distance has a direct impact on the delays between the departure and arrival airports [23]. Therefore, the distance matrix can select as the spatial weight matrix. The formula is as follows:

$$w_{ij} = \sqrt{\sum_{i=1, j=1}^{n} \left(x_i - x_j\right)^2}. \tag{9}$$

After reciprocal distance processing, geographic distance matrix is established and standardized processing is completed as formula (10) and formula (11):

$$w_{ij} = \left\{ \begin{array}{ccc} w_{11} & \cdots & w_{1n} \\ \vdots & \cdots & \vdots \\ w_{n1} & \cdots & w_{nn} \end{array} \right\}, \tag{10}$$

$$w_{ij} = \frac{w_{ij} - \min w}{\max w - \min w}. \tag{11}$$

The space distance matrix of airport delay includes a set of airport pairs (that is, two-dimensional array). Therefore, given N airports in the Euclidean space, the distance matrix is a symmetric $N * N$ matrix with nonnegative real numbers as elements. In the spatial delay analysis, the greater [24] the distance is, the smaller the impact of airports on the delay will be.

3.3. Modeling. In order to analyze the correlation of the key factors in the airport delay Internet of Things, we established a spatial regression model for analysis. Spatial econometric regression models can be achieved in many forms [26], while spatial lag models and spatial error models are commonly used types. When airport delays have spatial grid effects, it is necessary to establish a feature delay model that includes spatial relationships. Then, the spatial lag model or the spatial error model can be applied based on spatial autocorrelation and spatial heterogeneity.

The flight delay impact function can be received based on the analysis on the determinants of airport delays. The following equation is the basic quantitative regression model:

$$\begin{aligned} TD_{it} = {} & \alpha + \beta_1 AAC_{it} + \beta_2 DFC_{it} + \beta_3 NDF_{it} + \beta_4 PTD_{it} + \beta_5 ATL_{it} \\ & + \beta_6 NHA_{it} + \beta_7 ACR_{it} + \beta_8 WEA_{it} + \mu_i + \lambda_t + \varepsilon_{it}. \end{aligned} \tag{12}$$

After determining the spatial correlation, in the analysis of the delay characteristics, except for the estimation of the OLS estimation, it is necessary to consider the spatial regression model of spatial effects. Therefore, with the base of the traditional regression model, extra spatial weight matrix

can be adopted. The spatial Dubin model can be transformed to a spatial lag model (SLM) and a spatial error model (SEM) by setting constraints (when $\omega = 0$), so as to establish the only spatial Durbin model, and the model can obtained as the following equation:

$$\begin{aligned} TD_{it} = {} & \alpha + \delta WTD_{it} + \beta_1 AAC_{it} + \beta_2 DFC_{it} + \beta_3 NDF_{it} + \beta_4 PTD_{it} + \beta_5 ATL_{it} + \beta_6 NHA_{it} + \beta_7 ACR_{it} \\ & + \beta_8 WEA_{it} + \theta_1 WAAC_{it} + \theta_2 WDFC_{it} + \theta_3 WNDF_{it} + \theta_4 WPTD_{it} + \theta_5 WATL_{it} + \theta_6 WNHA_{it} \\ & + \theta_7 WACR_{it} + \theta_8 WWEA_{it} + \mu_i + \lambda_t + \varepsilon_{it}. \end{aligned} \tag{13}$$

The above model can also be displayed as the following equation:

$$TD_{it} = \alpha + \delta \sum_{j=1}^{100} w_{ij} TD_{it} + \beta_1 AAC_{it} + \beta_2 DFC_{it} + \beta_3 NDF_{it} + \beta_4 PTD_{it} + \beta_5 ATL_{it} + \beta_6 NHA_{it}$$

$$+\beta_7 ACR_{it} + \beta_8 WEA_{it} + \theta_1 \sum_{j=1}^{100} w_{ij} AAC_{it} + \theta_2 \sum_{j=1}^{100} w_{ij} DFC_{it} + \theta_3 \sum_{j=1}^{100} w_{ij} NDF_{it}$$

$$+\theta_4 \sum_{j=1}^{100} w_{ij} PTD_{it} + \theta_5 \sum_{j=1}^{100} w_{ij} ATL_{it} + \theta_6 \sum_{j=1}^{100} w_{ij} NHA_{it} + \theta_7 \sum_{j=1}^{100} w_{ij} ACR_{it}$$

$$+\theta_8 \sum_{j=1}^{100} w_{ij} EA_{it} + \mu_i + \lambda_t + \varepsilon_{it},$$

(14)

where TD_{it} is the explanatory variable, WTD_{it} is the spatial lag term of the explanatory variable, δ is the spatial autoregressive coefficient, AAC_{it}, DFC_{it}, PTD_{it}, ATL_{it}, NHA_{it}, ACR_{it}, and WEA_{it} are the explanatory variables, $WAAC_{it}$, $WDFC_{it}$, $WPTD_{it}$, $WATL_{it}$, $WNHA_{it}$, $WACR_{it}$, and $WWEA_{it}$ are the spatial lag term of the explanatory variables, W is the 100×100-order spatial distance matrix, θ_n is the explanatory variable space lag term coefficient, β_n is the regression coefficient, $n = 1, 2, 3, 4, 5, 6, 7, 8$, μ_i means a spatial effect, λ_t means time effect, ε_{it} is random error term, $\varepsilon_{it} = \omega \sum_{j=1}^{100} W_{ij} \varepsilon_{it} + \phi_{it}$ is spatial autocorrelation coefficient of error term, and ϕ_{it} is error term of independently and identically distribution.

4. Results

According to the data of delayed IoT mining and combined with the spatial model, the results in Table 3 show that the goodness of fit R2 for the basic model is 43.6%, and the adjusted R2 is 40%. F value is 8.8. The model passes the 1% level of the significance test, reaching obvious DFC and PTD on level 0.01, NHA, NHA, and WEA on level 0.05, while AAC, PTD, ATL, and ACR have passed a 10% level of the significance test.

As for the spatial correlation with different factors among the airports, four levels can describe the interspecific association, including high, higher, lower, and low. By applying regression analysis to all relevant determinants in the airport IoT network considering the average delay time, the spatial distribution of airport delay determinants can maintain the pattern of "high delay in the East and South area" and "low delay in the north and west area." Only a few airports have a different distribution.

About airport delay IoT grid distribution, Figure 4(a) shows that high-high agglomeration airports in North China have PEK, NAY, and TSN. The Yangtze River Delta region is an another high-high agglomeration area for flight delays, mainly including SHA, PVG, NKG, CZX, HGH, and NGB. The high-low clusters are mainly distributed between CTU, CAN, KMG, CKG, XMN, SYX and LJG, LZH, WXN, TCZ, while HTN, HLH, YIN, HET, and other airports are low-low clusters.

Figures 4(b)–4(i) show the delay factor distribution from AAC to WEA under the intelligent sensing and wireless communication data analysis. With the delay aggregation graph, the influence degree on each

determinant of delay on the regional distribution is discrepant, except the hardest-hit areas located in Beijing, Shanghai, and Guangzhou. On the other hand, it is worth noting that critical delays have also occurred in Xizang. Combined with the correlation analysis of delay factors, the results present that the influencing factors do not play a decisive role in Xizang delay. Therefore, besides the influencing factors, the delay in Xizang may also be caused by airport operation failures.

For the determinants in the IoT structure of airport delay, the AAC aggregation map can express the distribution of airport flow and capacity delays. Traffic can delay distribution, concentrated in the Shanghai Pudong Airport and Shanghai Hongqiao Airport, as a high-high agglomeration area in the AAC variable. Beijing Capital International Airport, Guangzhou Baiyun Airport, and Shanghai's two airports have a high-high cluster (means the delays situation of airport and its nearby airports in the surrounding areas are both serious.) at the DFC, NDF, and PTD airports. Except for the metropolises and first-tier cities, airports worthy of attention include Yunnan Changshui airport, Inner Mongolia Baita airport, and Sanya Phoenix Airport. As a tourist city with high demand throughout the year, they have a high degree of agglomeration in NHA, ACR, WEA, and PTD.

Moran's I that calculated with the GeoDa software has displayed the 0.67 airport delay, indicating that the airport delay is not entirely random in spatial distribution with specific spatial correlation. Figure 5 is a scatter plot of Moran's I. This shows that the spatial big data mined from each flight path in the airport IoT has spatial correlation in the airport delay grid.

Table 4 shows the regression results for OLS, SLM, and SEM. Comparing the basic model and the spatial model, the Log L values of SLM and SEM are more significant, the AIC value and the SC value are smaller than OAS's AIC and SC, the spatial model is better than the basic model, and the fitting effect is also better than OLS. Therefore, the traditional regression model may have specific limitations in analyzing delay, which also implies the necessity of the spatial regression model. On the other hand, comparing SLM and SEM, it is found that the SLM has a more substantial Log L value, a more significant LR value, a smaller AIC value and an SC value, and a better SLM estimation effect. The result shows that the delay between China's airports has a strong proximity effect, while the spatial

TABLE 3: The results of OLS model with multivariable in airport delay IoT.

| Variable	Ordinary least square (OLS)			
	Coef	Std. err	z	P
AAC	−0.0333	0.0374	−0.8898	0.3759
DFC	0.2881	0.0903	3.1908	0.0020
NDF	0.2252	0.0636	3.5436	0.0006
PTD	−0.0640	0.04461	−1.4333	0.1552
ATL	0.2706	0.1812	1.4932	0.1388
NHA	−0.0419	0.0163	−2.5669	0.0119
ACR	0.0145	0.0295	0.4921	0.6239
WEA	0.1077	0.0517	2.0816	0.0402
Constant	−0.3461	0.9085	−0.38092	0.7042
	$R2$	$R2$ (adj.)	F-statistic	Prob (F)
	0.4359	0.3963	8.7897	$7.82369e − 009$
	LogL.	AIC	SC	
	19.8936	−21.7872	1.6594	
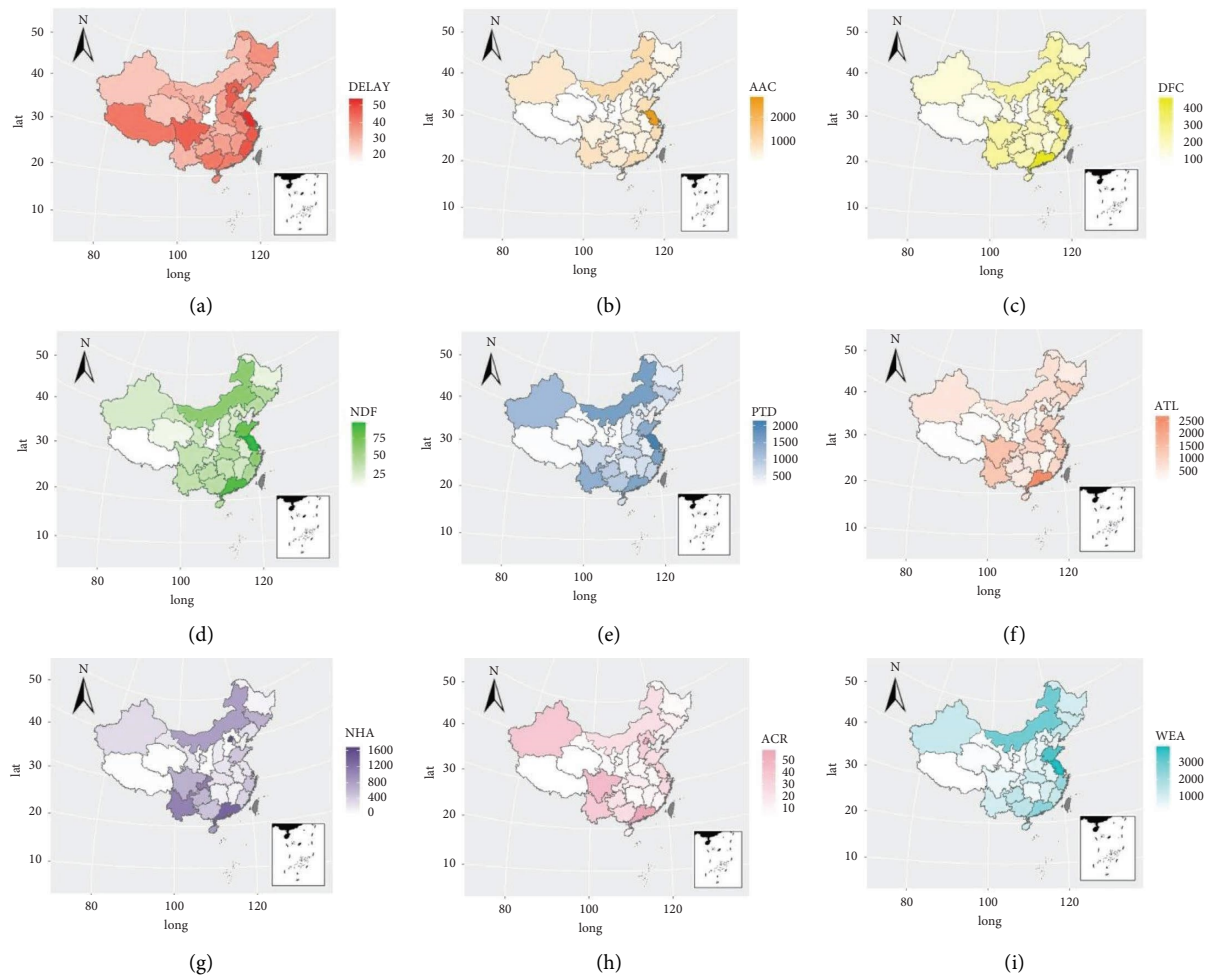
FIGURE 4: Aggregation map of an airport delay IoT system. (a) The distribution of delay grid. (b) The distribution of AAC. (c) The distribution of DFC. (d) The distribution of NDF. (e) The distribution of PTD. (f) The distribution of ATL. (g) The distribution of NHA. (h) The distribution of NHA. (i) The distribution of WEA).

heterogeneity of delays (errors) is relatively weak. In the SLM model, the airport delay has significant spatial effects with spatial correlation coefficient $\rho = 0.9863$, indicating that 100 airport delays have an extreme spatial dependence under the proximity effect in the airport grid.

As aforementioned, it is vital to provide some policy suggestions to reduce the aiport delay, as shown by conducting statitical analysis to the relevant intelligent flight sensing and communication data and combining the correlation degree of each determinant to airport delay in
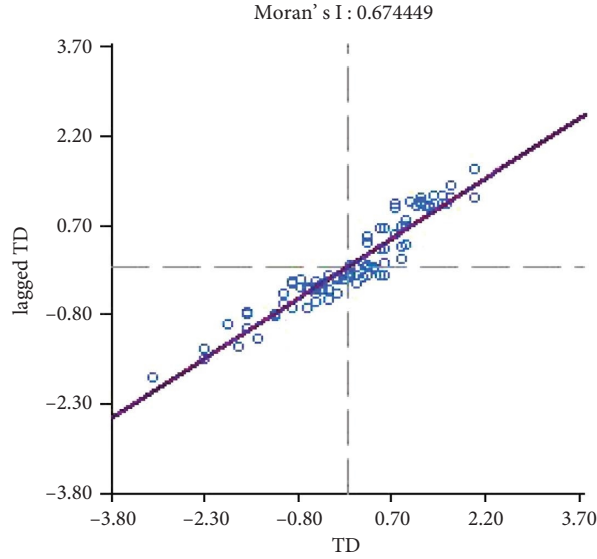
FIGURE 5: Moran's I scatter plot.

TABLE 4: Comparison of regression results between SLM and SEM.

Variable	SLM				SEM			
	Coef.	Std. err	z	P	Coef.	Std. err	z	P
AAC	−0.005	0.01367	−0.388	0.6978	0.0032	0.01586	0.2029	0.8392
DFC	0.0968	0.03364	2.8789	0.004	0.0849	0.03777	2.2485	0.0245
NDF	0.0616	0.0238	2.5882	0.0097	0.0247	0.02836	0.8706	0.384
PTD	−0.012	0.01716	−0.667	0.5045	−0.003	0.01913	−0.133	0.8944
ATL	0.0639	0.06781	0.9424	0.346	0.0403	0.07643	0.5273	0.598
NHA	−0.009	0.00609	−1.377	0.1687	−0.003	0.00682	−0.315	0.7524
ACR	−0.014	0.01079	−1.301	0.1934	−0.017	0.01209	−1.397	0.1623
WEA	0.02	0.01952	1.0252	0.3053	0.006	0.02285	0.2642	0.7917
WAAC	−0.031	0.07891	−0.393	0.6943	−0.019	0.08889	−0.215	0.8301
WDFC	−0.12	0.08469	−1.416	0.1569	−0.086	0.09464	−0.905	0.3655
WNDF	−0.049	0.12212	−0.395	0.6925	−0.022	0.13496	−0.165	0.8687
WPTD	0.0031	0.05811	0.0544	0.9566	−0.016	0.06491	−0.252	0.8012
WATL	0.0551	0.11145	0.4946	0.6209	0.004	0.12643	0.0318	0.9747
WNHA	0.0319	0.03677	0.8675	0.3857	0.0208	0.04078	0.51	0.61
WACR	0.0408	0.05215	0.7826	0.4338	0.0356	0.05909	0.602	0.5471
WWEA	0.0773	0.12081	0.6401	0.5221	0.0851	0.13443	0.6332	0.5266
Constant	−1.088	0.37027	−2.938	0.0033	2.3556	1.12365	2.0964	0.036
ρ	0.9863	0.00931	105.99	0	—	—	—	—
λ	—	—	—	—	0.9918	0.00564	175.99	0
R^2		0.920964				0.89792		
Log L		113.067				99.725176		
AIC		−190.134				−165.45		
SC		−143.241				−121.162		

the spatial dimension. Due to the different scales, systems, and natural conditions, the policy should consider the diversity of the environmental conditions among the airports especially in the rapid development stage of IoT technology. The number of direct flights and new flights has a more significant impact on airport delays. However, in recent years, the quantity demand and planning of new airports and flight routes would increase by verging to multifold levels. The overall optimization can fundamentally reduce airport flight delays and control the economic loss while meeting the transportation demand.

Take-off and landing conditions in the delay IoT system are also important factors that affect the flight delay. Airport congestion mainly comes from aircraft operation flow and flight capacity. Combined with the spatial distance, taking the delay impact of regional influencing factors into account as airport take-off and landing queue structure adjustment and evacuation may bring unexpected benefits to solve the delay issues. In the airport IoT network, flight delay also needs to consider a transmission delay. The congestion of airspace capacity would also affect the operational structure of the airport. On the other hand, although the passenger

throughput and the allocation of heavy aircraft will affect the delay of the airport, they are negatively related to the delay in the spatial. Therefore, when considering the allocation of passenger flow and aircraft types, airlines only need to consider its impact on the delay of a single airport.

Although the occurrence of flight delay is irresistible and the recovery of delay is uncontrollable when special weather conditions occur, the weather information system is a decisive factor with particular regional characteristics. According to the agglomeration distribution, in this case, Sanya, Haikou, and Shenzhen are cities which greatly affected by the weather. At the same time, due to the interaction of spatial distance between cities, marine climate characteristics, and the weather among airports, there is a strong correlation between the three airports in weather-induced delays. The weather conditions shall also be considered when the airport conducts a route schedule. However, from the overall point of view, compared with other determinants, there are few and concentrated areas strongly affected by weather factors. So, it is relatively easy to improve the delay in the high concentration area caused by weather factors.

5. Conclusion

The Internet of Things, big data, artificial intelligence, 5G, and other new technologies are still in continuous improvement. With these technologies, flight delay has formed a framework of delay mining technology systems in smart airport grids with advanced technology, open data fusion, security, and reliability. After the data set of airports, aircraft, flights, and passengers is collected through the IoT technology, this paper firstly utilized spatial autocorrelation to process multivariables-delay determinants (using intelligent flight location and communication data) and analyzed the spatial distribution characteristics of the airport delays grid in China. Basic regression models and geographically weighted regression models were then used to study the driving factors and regional differences in an airport delays information system. The results demonstrated the validity of the spatial econometric regression model. Secondly, the spatial aggregation characteristics of China's airport delays were high aggregation and low accumulation, while DFC, NDF, and WEA positively correlated with the delay time. At the same time, the results of the geographically weighted regression model revealed that different spatial differences between multiple factors would lead to diverse effects of civil aviation delays. The spatial regression model can more systematically and intuitively understand various determinants in the airport system in different regions. This method can also be applied to other relevant studies, for example, time-space analysis of delay factors, or apply the findings to delays in the study of systemic grid structure propagation and delay assessment.

The delay space analysis of the aircraft execution process is an effective measure to ensure the operation of flights. The results of delay distribution in the airport grid are directly related to the adjustment and control of flight operations. Based on the analysis of multiple determinants, big data and

multiattribute data excavated by the Internet of Things application development techniques are used to conduct spatial assessment modeling. From the perspective of flight trajectory, weather, and passenger demand learning, it is significant to explore the spatial distribution level, delay diffusion level, and determinants of delay during flight execution. Furthermore, the results of spatial delay analysis can help the airport IoT technology to reallocate flight operation support equipment and facilities according to the distribution of abnormal areas involved in a flight delay.

Data Availability

The related data used to support the findings of this study were supplied by Xiushan Jiang under license and so cannot be made freely available. Requests for access to these data should be made to Xiushan Jiang, xshjiang@bjtu.edu.cn.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] Y. Chen, J. Yu, L. Li, J. Zhou, S. B. Tsai, and Q. Chen, "An empirical study of the impact of the air transportation industry energy conservation and emission reduction projects on the local economy in China," *International Journal of Environmental Research and Public Health*, vol. 15, no. 4, p. 812, 2018.

[2] F. Aviation Administration, "The economic impact of civil aviation on the U.S. Economy," *FAA Economic Impact Report 2009*, vol. 77, p. 188, 2016.

[3] F. Aviation Administration, "Airport capacity and delay," *FAA Advisory Circular*, 2016, https://www.faa.gov/documentLibrary/media/Advisory_Circular/150_5060_5. pdf.

[4] University of Westminster, "European airline delay cost reference values," *Updated and extended values Version 4.1*, 2015, https://www.eurocontrol.int/sites/default/files/publication.

[5] Acca report, "Statistical Bulletin on the development of civil aviation industry in 2017," *Civil Aviation Administration of China*, 2018, http://www.caac.gov.cn.

[6] Bts report, "Airline on-time performance and causes of flight delays," *Beurea of Trasportation*, 2017, https://www.bts.gov/explore-topics-and-geography.

[7] R. Beatty, R. Hsu, L. Berry, and J. Rome, "Preliminary evaluation of flight delay propagation through an airline schedule," *Air Traffic Control Quarterly*, vol. 7, no. 4, pp. 259–270, 1999.

[8] Y. J. Wang, Y. K. Cao, C. P. Zhu et al., "Characterizing departure delays of flights in passenger aviation network of United States," 2017, http://arxiv.org/org/1701.05556.

[9] A. Alghadeir and H. A. Sakran, "Smart airport architecture using internet of things," *International Journal of Innovative Research in Computer Science & Technology (IJIRCST) ISSN*, vol. 34, pp. 2347–5552, 2016.

[10] H. Ning, S. Hu, W. He, Q. Xu, H. Liu, and W. Chen, "nID-based Internet of Things and its application in airport aviation risk management," *Chinese Journal of Electronics*, vol. 21, no. 2, pp. 209–214, 2012.

[11] Q. Qi and Z. Pan, "Internet of things, internet, big data and airport services make smart airport based on o2o and humanism," in *Proceedings of the 2018 International Conference on Mechanical, Electronic, Control and Automation Engineering (MECAE 2018)*, Atlantis Press, Philippines, March 2018.

[12] M. Hansen and C. Y. Hsiao, "Going south? econometric analysis of US airline flight delays from 2000 to 2004," *Transportation Research Record*, vol. 1915, no. 1, pp. 85–94, 2005.

[13] B. Zou, M. Elke, M. Hansen, and N. Kafle, "Evaluating air carrier fuel efficiency in the US airline industry," *Transportation Research Part A: Policy and Practice*, vol. 59, pp. 306–330, 2014.

[14] R. Duran-Fernandez and G. Santos, "A regional model of road accessibility in Mexico: accessibility surfaces and robustness analysis," *Research in Transportation Economics*, vol. 46, no. 20, pp. 55–69, 2014.

[15] A. Lall, "Delays in the New York city metroplex," *Transportation Research Part A: Policy and Practice*, vol. 114, pp. 139–153, 2018.

[16] Acca report, "Normal statistics measures for civil aviation flights (draft for comments)," *Civil Aviation Administration of China*, 2018, http://www.caac.gov.cn.

[17] P. Baumgarten, R. Malina, and A. Lange, "The impact of hubbing concentration on flight delays within airline networks: an empirical analysis of the US domestic market," *Transportation Research Part E: Logistics and Transportation Review*, vol. 66, pp. 103–114, 2014.

[18] Á. Rodríguez-Sanz and L. R. Andrada, "Managing airport capacity and demand: an economic approach," *IOP Conference Series: Materials Science and Engineering*, vol. 1226, no. 1, Article ID 012024, 2022.

[19] I. Simaiakis, "Analysis, modeling and control of the airport departure process," *Doctoral dissertation, Massachusetts Institute of Technology*, vol. 66, pp. 223–245, 2013, http://dspace.mit.edu/handle/1721.1/79342.

[20] W. B. Du, M. Y. Zhang, Y. Zhang, X. B. Cao, and J. Zhang, "Delay causality network in air transport systems," *Transportation Research Part E: Logistics and Transportation Review*, vol. 118, pp. 466–476, 2018.

[21] D. Markovic, T. Hauf, P. Röhner, and U. Spehr, "A statistical study of the weather impact on punctuality at Frankfurt Airport," *Meteorological Applications*, vol. 15, no. 2, pp. 293–303, 2008.

[22] M. Schultz, S. Lorenz, R. Schmitz, and L. Delgado, "Weather impact on airport performance," *Aerospace*, vol. 5, no. 4, p. 109, 2018.

[23] B. Zou and M. Hansen, "Flight delay impact on airfare and flight frequency: a comprehensive assessment," *Transportation Research Part E: Logistics and Transportation Review*, vol. 69, pp. 54–74, 2014.

[24] S. Reitmann, S. Alam, and M. Schultz, "Advanced quantification of weather impact on air traffic management," *ATM Seminar*, vol. 44, pp. 554–567, 2019.

[25] J. P. LeSage and R. K. Pace, "Spatial econometric modeling of origin-destination flows," *Journal of Regional Science*, vol. 48, no. 5, pp. 941–967, 2008.

[26] S. S. Allan, J. A. Beesley, J. E. Evans, and S. G. Gaddy, "Analysis of delay causality at newark international airport," *USA/Europe Air Traffic Management R&D Seminar. 4th USA/Europe Air Traffic Management R&D Seminar*, vol. 55, p. 665, 2001.