

Article

Evaluating Prediction Models for Airport Passenger Throughput Using a Hybrid Method

Bin Chen ^{1,2} , Xing Zhao ^{1,*} and Jin Wu ^{1,*}¹ Civil Aviation College, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China² China Civil Aviation Engineering Consulting Co., Ltd., Beijing 100621, China

* Correspondence: xingzhao@nuaa.edu.cn (X.Z.); wujin@nuaa.edu.cn (J.W.)

Abstract: This paper proposes a hybrid evaluation method to assess the prediction models for airport passenger throughput (APT). By analyzing two hundred three airports in China, five types of models are evaluated to study the applicability to different airports with various airport passenger throughput and developing conditions. The models were fitted using the historical data before 2014 and were verified by using the data from 2015–2019. The evaluating results show that the models employed for evaluating perform well in general except that there are insufficient historical data for modelling, or the APT of the airports changes abruptly owing to expansion, relocation or other kinds of external forces such as earthquakes. The more the APT of an airport is, the more suitable the models are for the airport. Particularly, there is no direct relation between the complexity and the predicting accuracy of the models. If the parameters of the models are properly set, time series models, causal models, market share methods and analogy-based methods can be utilized to predict the APT of 88% of studied airports effectively.

Keywords: airport; throughput prediction; adaptation analysis; ex-post evaluation



Citation: Chen, B.; Zhao, X.; Wu, J. Evaluating Prediction Models for Airport Passenger Throughput Using a Hybrid Method. *Appl. Sci.* **2023**, *13*, 2384. <https://doi.org/10.3390/app13042384>

Academic Editor: Vicente Julian Inglada

Received: 8 December 2022

Revised: 6 February 2023

Accepted: 9 February 2023

Published: 13 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

An accurately predicting result of passenger throughput for an airport is one of the most important factors for its development decision on construction and expansion. A larger predicted result will inevitably lead to a waste of resources such as untimely land occupations, idle facility constructions and lower returns on investments. When the prediction result is too small, the crowded airport terminal may bring about poor travel experiences as well as higher security risks. Therefore, conducting accurate predictions will effectively avoid economic losses caused by inappropriate decision-making during airport development [1].

The task for more accurate predictions has been studied since early ages. Predictions can be conducted both qualitatively by using methods such as executive judgement, market research and Delphi techniques, and quantitatively by using methods including time series forecasting [2,3] and causal models [4–6]. To promote the precision of prediction and the adaptivity for different types of airports, hybrid approaches integrating multiple models are proposed. Feng et al. [7] proposed a hybrid method assembling variational mode decomposition, autoregressive moving average model and kernel extreme learning machine, which showed a better result than the single-model methods. Xie et al. [8] proposed two hybrid approaches based on seasonal decomposition and the least squares support vector regression model for short-term forecasting of air passengers. Their empirical analysis showed that the proposed hybrid approaches were better than other time series models. Currently, artificial intelligence is performing stronger ability on different prediction scenes. Artificial intelligence methods including support vector regression (SVR), Monte Carlo simulation, decision-making tree and deep learning are widely used for airport passenger throughput (APT) prediction [9,10]. Scarpel [11] employed an integrated mixture of local

expert models to forecast air passengers at São Paulo International Airport. The model was validated using out-of-sample data, and the accuracy of the generated predictions proved to be satisfactory. Zhou et al. [12] proposed a grey seasonal least square support vector regression, and it was highly recommended for addressing issues with periodic and nonlinear features. Wong et al. [13] combined a Markovian model with a grey model and they found that a fuzzy Markovian model showed better performance on the observations with trends and intercepts. Meanwhile, big data are providing possibilities for real-time and more precise APT prediction, especially for specific short-time prediction scenes. Liang et al. [14] related the data from search engines to short-time passenger demands and proposed a novel decomposition ensemble model to discuss the role of Internet search data in air passenger demand forecasting. The decomposition ensemble model obtained more accurate and reliable prediction results than the benchmark models. Li et al. [15] proposed a method for separating the two forces of COVID-19 and evaluating the respective impact on demand, dividing passengers into different segments based on passenger characteristics, simulating different scenes, and predicting demand for each passenger segment in each scene. Barczak et al. [16] used a time series model to study the difference between the demand that was observed during the pandemic, and the demand that was forecast based on the pre-pandemic trend. All models suggest that demand would have increased further without COVID-19.

Since the development of airport passenger throughput is related to many stochastic factors, it is difficult to observe the causal links between APT and potential factors. In addition, the accuracies of different prediction methods cannot be compared without a clear and definite prediction scene. Therefore, studies have been conducted to judge the performance of different prediction methods. Matthew [17] proposed a synthetic method to assess the forecasting quality of time series methods and time series models with econometric independent explanatory variables on Miami International Airport and Frankfurt Airport, which indicated that simple models with few independent variables performed as well as more complicated and costly models and that external factors had a pronounced effect on air-travel demand. Maldonado [18] and Mierzejewski [19] evaluated the 5-year, 10-year and 15-year predicting results of 22 airports distributed in the USA, and the evaluations showed that the standard deviation of the predicting results was distributed with very large errors from 30% to 69%.

According to the studies on ex-post evaluation of APT prediction, a reasonable framework should be established priorly to evaluating the performances of different forecasting methods [20–22]. The performances of different forecasting methods are mainly affected by the models and applicable scenes. Influenced by stochastic factors, the models for APT prediction are intuitively inaccurate with plentiful assumptions. The assumptions determined that different models are suitable for certain prediction scenes. A hybrid evaluating method is then proposed in this paper, adjusting to existing prediction methods, and considering multiple prediction scenes.

Aiming at evaluating the performance of various methods for APT prediction, the APT of 203 airports in China are studied in this paper. A proposed evaluation method for five common prediction models is introduced. The hypothesis and the evaluating scenes of the proposed evaluation method are put forward. The evaluating method is conducted on 203 airports in China, and the evaluating results of different forecasting methods are obtained. An accurately predicting result of passenger throughput for an airport is one of the most important factors for its development decision on construction and expansion. This paper analyzed the airport passenger throughput (APT) of 203 airports in China, evaluating the performance of various methods for APT prediction and introducing a proposed evaluation method for five common prediction models. It studied the applicability of the five models to different airports with various developing conditions.

2. The Proposed Hybrid Evaluating Method

Various approaches have been proposed to forecast the APT of airports. The modelling methods can be divided into four categories including time series models, causal models, artificial intelligent models and hybrid models. In addition, market share methods and analogy-based methods are also widely used for forecasting the construction scale in the field of engineering. In this paper, time series models, causal models, artificial intelligent models, market share methods and analogy-based methods are mainly studied. Taking the five methods as a whole, a hybrid evaluating method is built to see if the airport passenger throughput can be predicted. When the hybrid evaluating method of the five common modelling models shows high prediction accuracy, the best prediction result can be obtained using one of the appropriate models as long as the parameter selection is reasonable and the method selection is appropriate. However, other unique methods should be considered when the hybrid evaluating method shows bad prediction accuracy.

2.1. Models

2.1.1. Time Series Models

Time series models regard the systems for prediction as black boxes, without considering the factors which influence the results. Instead, the models are established by fitting several historical data. Trend extrapolation models, exponential smoothing models, grey models, autoregressive models and moving average models are typical time series models, which have been widely used in nearly every prediction scene. Considering the impact of specific events, complex methods may perform better. For example, Djakaria et al. [23] predicted the passenger demand of Djalaluddin Gorontalo Airport using a multiplicative of Holt-Winters exponential smoothing. Elena et al. [24] compared the performance of various models including linear trend, quadratic trend, exponential trend, linear exponential smoothing (Holt’s Model), and autoregressive integrated moving average models on APT prediction, and the results showed that linear exponential smoothing model performed best facing the impact of COVID-19, with a level of reliability of 95%. Concerning the large number of research samples, linear trend, quadratic trend, cubic trend, power trend, exponential trend, logarithmic trend, quadratic exponential smoothing, cubic exponential smoothing and grey GM (1,1) models are considered in this paper.

To define the models for evaluation, an APT observation sequence $x_{0:k} = \{x_0, x_1, \dots, x_k\}$ is proposed, where k denotes a time metric.

The parameters of the linear trend model, the quadratic trend model, the cubic trend model, the power trend model, the exponential trend model and the logarithmic trend model by using the least squares method and the models can be described as:

$$\begin{aligned}
 \text{Linear trend model : } \hat{x}_{lf} &= a_{lf}t + b_{lf} \\
 \text{Quadratic trend model : } \hat{x}_{qf} &= a_{qf}t^2 + b_{qf}t + c_{qf} \\
 \text{Cubic trend model : } \hat{x}_{cf} &= a_{cf}t^3 + b_{cf}t^2 + c_{cf}t + d_{cf} \\
 \text{Power trend model : } \hat{x}_{pf} &= a_{pf}e^{b_{pf}x} \\
 \text{Exponential trend model : } \hat{x}_{pf} &= a_{pf}x^{b_{pf}} \\
 \text{Logarithmic trend model : } \hat{x}_{\log f} &= a_{\log f} \log t + b_{\log f}.
 \end{aligned} \tag{1}$$

The quadratic exponential smoothing model is based on a single exponential smoothing model, and the single exponential smoothing model can be obtained by:

$$S_{t+1}^{es1} = \alpha_1 x_t + (1 - \alpha_1) S_t^{es1} \tag{2}$$

where α_1 is the smoothing coefficient of the model, S_t^{es1} is the single smoothing result at time t and S_{t+1}^{es1} is the single smoothing result at time $t + 1$. When $t = 1$, S_0^{es1} is set to be the value of x_0 . Then, the quadratic exponential smoothing model can be deduced as:

$$S_{t+1}^{es2} = \alpha_2 S_t^{es1} + (1 - \alpha_2) S_t^{es2} \tag{3}$$

where α_2 is the smoothing coefficient of the model, S_t^{es1} is the single smoothing result at time t , S_t^{es2} is the quadratic smoothing result at time t and S_{t+1}^{es2} is quadratic the smoothing result at time $t + 1$. When $t = 1$, S_0^{es2} is set to be the value of S_0^{es1} . The predicted result of APT \hat{x}_{t+T}^{es2} at time $\hat{x}_{t+T}^{es2} = a_{2t} + b_{2t} * Tt + T$ can be calculated by:

$$\hat{x}_{t+T}^{es2} = a_{2t} + b_{2t} * T \tag{4}$$

$$a_t^{es2} = 2S_t^{es1} - S_t^{es2} \tag{5}$$

$$b_t^{es2} = (\alpha_2 / (1 - \alpha_2)) * (S_t^{es1} - S_t^{es2}) \tag{6}$$

where a_t^{es2} and b_t^{es2} are the parameters which can be calculated by Equations (5) and (6).

Similarly, the cubic exponential smoothing model is based on the quadratic exponential smoothing model, which is defined as:

$$S_{t+1}^{es3} = \alpha_3 S_t^{es2} + (1 - \alpha_3) S_t^{es3} \tag{7}$$

where α_3 is the smoothing coefficient of the model, S_t^{es3} is the quadratic smoothing result at time t and S_{t+1}^{es3} is quadratic the smoothing result at time $t + 1$. The predicted result of APT \hat{x}_{t+T}^{es3} at time $t + T$ can be calculated by:

$$\hat{x}_{t+T}^{es3} = a_t^{es3} + b_t^{es3} * T + c_t^{es3} T^2 \tag{8}$$

$$a_t^{es3} = 3S_t^{es1} - 3S_t^{es2} + S_t^{es3} \tag{9}$$

$$b_t^{es3} = \frac{\alpha_3}{2(1 - \alpha_3)^2} [(6 - 5\alpha_3) S_t^{es1} - 2(5 - 4\alpha_3) S_t^{es2} + (4 - 3\alpha_3) S_t^{es3}] \tag{10}$$

$$c_t^{es3} = \frac{\alpha_3^2}{2(1 - \alpha_3)^2} [S_t^{es1} - 2S_t^{es2} + S_t^{es3}] \tag{11}$$

where a_t^{es3} , b_t^{es3} and c_t^{es3} are the parameters which can be calculated by Equations (9)–(11).

For the grey model GM (1,1), the raw sequence $X^{(0)} = \{x_0, x_1, \dots, x_k\}$ can be accumulated by $X^{(1)} = \sum_{i=0}^k x_i, k = 0, 1, 2, \dots$ to obtain the accumulated sequence. Then, the accumulated sequence can be used to fit $\frac{dx_k^{(1)}}{dt} + a^{gr} x_k^{(1)} = b^{gr}$ by:

$$\begin{bmatrix} a^{gr} \\ b^{gr} \end{bmatrix} = (\mathbf{B}^T \cdot \mathbf{B})^{-1} \cdot \mathbf{B}^T \cdot \mathbf{Y} \tag{12}$$

where

$$\mathbf{B} = \begin{bmatrix} -0.5 \times (x_1^{(1)} + x_0^{(1)}) & 1 \\ -0.5 \times (x_2^{(1)} + x_1^{(1)}) & 1 \\ \vdots & \vdots \\ -0.5 \times (x_k^{(1)} + x_{k-1}^{(1)}) & 1 \end{bmatrix}, \tag{13}$$

And

$$\mathbf{Y} = [x_1^{(1)} \quad x_2^{(1)} \quad \dots \quad x_k^{(1)}]^T. \tag{14}$$

The predicted value of the accumulated sequence can be primarily obtained by:

$$\hat{x}^{(1)}(t) = (x^{(0)}(1) - \frac{b^{gr}}{a^{gr}}) e^{-a^{gr}t} + \frac{b^{gr}}{a^{gr}}, t = 1, 2, \dots, k \tag{15}$$

and the predicted result of APT $\hat{x}^{(0)}(t + 1)$ at time $t + 1$ can be iteratively calculated by:

$$\hat{x}^{(0)}(t + 1) = \hat{x}^{(1)}(t + 1) - \hat{x}^{(1)}(t) = (1 - e^{-a^{gr}t})(x^{(0)}(1) - \frac{b^{gr}}{a^{gr}}) e^{-a^{gr}t}, t = 1, 2, \dots, k \tag{16}$$

2.1.2. Causal Models

Causal models establish causal relationships between independent variables. Typical causal models include regression models and elastic coefficient models [25]. For regression models, a unary linear regression model, a multiple linear regression, a stepwise regression model, a hybrid regression model, an elastic coefficient model and a proposed elastic-like scale model are considered.

To define the models for evaluation, consider there are M correlated variables, the variables can be collected as $W(t) = \{W_1(t), W_2(t), \dots, W_m(t), \dots, W_M(t)\}$, where $W_m(t)$ is the value of m th variable at time point t . For APT prediction, the variables are usually the economic indicators of the cities where the airports are located in.

The unary linear regression model can be by fitting:

$$\hat{x}_{ulr}(t) = a_1 W_m(t) + e_1, m \in 1 \dots M \tag{17}$$

using the least squares method. Furthermore, the multiple linear regression can be similarly described as:

$$\hat{x}_{plr}(t) = c_1 W_1(t) + c_2 W_2(t) + \dots + c_p W_p(t) + e_1 + \dots + e_p \tag{18}$$

By combing the regression models with the time series models, $W_m(t + T)$ can be calculated first and then, $\hat{x}_{ulr}(t + T)$ and $\hat{x}_{plr}(t + T)$ can be obtained.

The stepwise regression model introduces correlated variables step-by-step into the regression model until the model reflects the relationship significantly. The common methods to determine the variables include the forward method and backward method. Taking the forward method as an example, the modelling process is shown as follows.

Step 1: Establishing the unary linear regression models between APT and each variable $W_1, W_2, \dots, W_m, \dots, W_M$, which can be obtained by Equation (17). Calculating the F inspection values of each regression coefficient of the models. The F inspection values can be denoted as $F_1^{(1)}, \dots, F_M^{(1)}$, and the maximum value can be obtained by $F_{m_1}^{(1)} = \max(F_1^{(1)}, \dots, F_M^{(1)})$. For a given significance level α , the critical value is denoted as $F^{(1)}$. If it satisfies that $F_{m_1}^{(1)} \geq F^{(1)}$, the variable W_{m_1} is selected as the regression variable and collected into set I_1 .

Step 2: Establishing the binary sets of W_{m_1} with the other variables, which can be denoted as $\{W_{m_1}, W_1\}, \dots, \{W_{m_1}, W_{m_1-1}\}, \{W_{m_1}, W_{m_1+1}\}, \dots, \{W_{m_1}, W_M\}$. Building the binary linear regression models between APT and the established binary sets. Calculating the F inspection values of each regression coefficient of the models. The F inspection values can be denoted as $F_m^{(2)} (m \notin I_1)$, and the maximum value can be obtained by $F_{m_2}^{(2)} = \max(F_1^{(2)}, \dots, F_{i_1-1}^{(2)}, F_{i_1+1}^{(2)}, F_M^{(2)})$. For a given significance level α , the critical value is denoted as $F^{(2)}$. If it satisfies that $F_{m_2}^{(2)} \geq F^{(2)}$, the variable W_{m_2} is selected as the regression variable and collected into set I_1 . Otherwise, the process is stopped.

Step 3: Similarly, considering establishing multiple linear regression models. The variables are selected into I_1 , until the process is stopped.

A hybrid regression causal model is established by using the correlation analysis method and the unary linear regression model. The correlation between the APT and the variables is first analyzed. The most correlated variable is selected to build the unary linear regression model with APT. The correlation coefficient can be obtained by:

$$\rho(x, W_m) = \frac{\sum_{t=1}^k (x(t) - \bar{x})(W_m(t) - \bar{W}_m)}{\sqrt{\sum_{t=1}^k (x(t) - \bar{x})^2} \sqrt{\sum_{t=1}^k (W_m(t) - \bar{W}_m)^2}} \tag{19}$$

where $\rho(x, W_m)$ is the correlation coefficient of the APT x and the variable W_m , $x(t)$ is the APT of the studied airport at time t , \bar{x} is the mean value of APT, $W_m(t)$ is the value of W_m at time t and \bar{W}_m is the mean value of W_m .

The elastic coefficient models are indirect methods to forecast the results by fitting the correlated factors. An elastic coefficient model can be defined as:

$$E_s = \frac{p'}{q'} \tag{20}$$

$$p = E_s * q \tag{21}$$

$$\hat{x}_{ec}(t + T) = (1 + p)^T \tag{22}$$

where E_s is the elastic coefficient, T is the target time point, q' is the growth rate of a correlated variable before time t , p' is the growth rate of the APT of an airport before time t , p is the growth rate of the APT of an airport before time T , q is the growth rate of a correlated variable before time T , $\hat{x}_{ec}(t + T)$ is the predicting result.

Zhang et al. [26] analyzed the causal relationship between air transport and economic growth, and the results showed the relationship was bi-directional, especially for the underdeveloped area. For the developed area, air transport only showed a positive effect on economic growth. The relationship can be reflected by the relative value of the APT and that of economic indicators, which can be represented by:

$$R_{AI} = \frac{APT}{indicator}. \tag{23}$$

Regarding the relative value R_{AI} as the elastic coefficient, an elastic-like scale model can be established. By fitting the indicators with trend extrapolation models, the predicted values of the indicators can be obtained and then, APT can be predicted.

2.1.3. Market Share Methods

The market share methods predict the value by forecasting the value of market size and the proportion of the studied APT with respect to the market size. Supposing that at time t , the APT of an airport is $x(t)$ and the market size is $q(t)$. Then, the market share of the airport is:

$$m(t) = \frac{x(t)}{q(t)} \tag{24}$$

When it comes to time $t + T$, supposing the market size is $q(t + T)$ and at the same time, the market share becomes $m(t + T)$, then the APT at time $t + T$ can be predicted by:

$$\hat{x}_{ms}(t + T) = q(t + T) * m(t + T) \tag{25}$$

2.1.4. Analogy-Based Method

The analogy-based method was first used for economic business forecasting in the 1920s [27], a forecasting process was proposed so that the experts can use the process to conduct analogy. Solvoll et al. [1] carried out verification on an airport in Norway to compare the performances of elastic models and analogy-based methods, and found that under particular circumstances, analogy-based methods performed better. The employed process is shown below.

Supposing that x_{it} is the APT of the studied airport at time t , x_{jt} is the APT of the analogical airport at time t . The predicted result $\hat{x}_{i(t+T)}$ can be obtained by the following steps.

Step 1: Determining the target predicting time T .

Step 2: Determining the conditions for analogy. For the APT prediction, the analogical condition can be determined by referring to the airport with a similar APT and setting the allowed error θ . The analogical condition can be described as $x_{jt} * (1 - \theta) \leq x_{it} \leq x_{jt} * (1 + \theta)$.

Step 3: Filtering the airports which satisfy the analogical condition from a database and all the J satisfied airports are collected as a set D .

Step 4: Obtaining the airports' APT data from D as a collection $X_{j(t+T)}$.

Step 5: The predicted result can be calculated by:

$$\hat{x}_{i(t+T)} = \frac{\sum_{j=1}^J x_{j(t+T)}}{J} \tag{26}$$

2.1.5. Artificial Intelligent Model

For nonlinear problems, back-propagation neural networks are often utilized to build the models [28–30]. In this paper, a typical back-propagation neural network is employed to conduct the evaluation. The typical back-propagation neural network consists of a three-layer framework, i.e., input–hidden–output layer, as is shown in Figure 1.

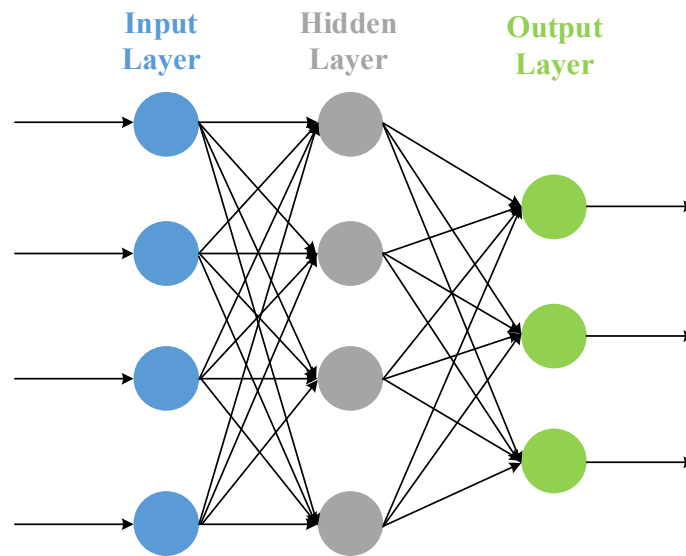


Figure 1. The structure of the typical back-propagation neural network.

The structure of a neuron is shown in Figure 2. For each neuron, the relationship between the input $\{x_1, \dots, x_i, \dots, x_n\}$ and the output y can be calculated as:

$$y = f(\sum w_i x_i) \tag{27}$$

where w is the weight for each input, Σ is to sum up all the inputs and f is the activation function.

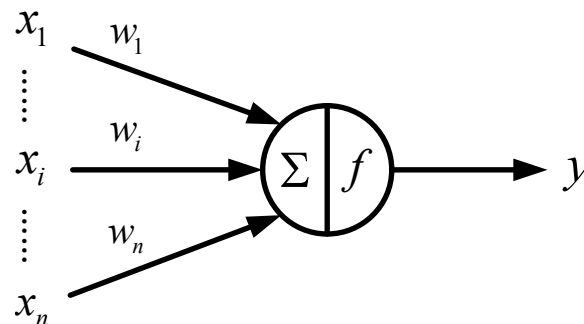


Figure 2. The structure of an artificial neuron.

2.2. Error Measures

Indicators used for evaluating the performances of APT prediction methods usually include mean error, mean square error (MSE), mean absolute error, mean percentage error and mean percentage absolute error [20,31,32]. For indicators such as mean percentage error and mean absolute percentage error (MAPE), when the value of the indicator reaches a threshold, the predicted results are believed as effective results. For the indicators such as MSE and mean absolute error, the smaller the absolute values of the indicators are, the better the predicted results are believed [33]. Particularly, for APT prediction, the predicted result of the APT in the target year draws great attention, which directly determines the next expansion time of the airport. In this paper, the MAPE is employed as the indicator to verify the adaption of different models for the target year prediction. The MAPE can be obtained by:

$$e_{it} = \left| \frac{Y_{it} - y_{it}}{y_{it}} \right| * 100\%, \quad (28)$$

where e_{it} is the APT predicting error of the i th airport in t th year, Y_{it} is the APT predicting the result of the i th airport in t th year and y_{it} is the APT observation of the i th airport in t th year.

The acceptable error of the predicted result can be defined as:

$$e_{it} \leq \varepsilon, \quad (29)$$

where ε is the maximum MAPE that can be accepted. Usually, when $e_{it} \leq 10\%$, the models are regarded as good models. When $e_{it} > 20\%$, the models are believed to be failure models.

The Civil Aviation Administration of China (CAAC) divides airports into three levels according to the APT of each airport. The distribution of 203 studied airports is shown in Table 1. Since the APT fluctuations of small airports and new airports are usually larger than that of the large airports and the airports under stable operation, ε is set to be 10%, 15% and 20% to study the available models for different types of airports. Four evaluation criteria are set in this paper to see the change of the number of predictable airports after changing the maximum allowable prediction error percentage.

Table 1. The distribution of 203 studied airports.

Level	Number of Airports	Percentage
APT less than 2 million	39	19%
APT between 2 million and 10 million	34	17%
APT more than 10 million	130	64%

3. Data and Hypothesis

For a comprehensive assessment of models for airports under different development conditions, 203 airports in China are studied, which covers various influence factors such as conditions of economic development, conditions of industrial structure and conditions of infrastructure. All the 203 airports have been under operation for at least 5 years, which avoids the impact of APT fluctuation during earlier times. Due to the breaking out of COVID-19, the APT of each airport decreased sharply in 2020 and 2021 [34]. The APT of all the airports in China in 2020 was only 63.3% in comparison to that in 2019, and such conditions will inevitably lead to the failures of all the models. Since the completion times of each airport are diverse from each other, the observations of APT before 2014 are used to fit the initial models. For each airport, the number of observation data for fitting is at least 15, which ensures fitting accuracy. The remaining observations, 2015–2019, are used to evaluate the forecasting performance of the models.

The data for the research can be divided into two types. One is the historical APT observations, which can be obtained from the Civil Aviation Administration of China (CAAC, <http://www.caac.gov.cn/en/SY/> (accessed on 22 August 2022)). The other one is the indicators of economic development, which contains gross domestic product (GDP), GDP

of primary industry, GDP of secondary industry, GDP of tertiary industry, resident population, year-end population, urbanization rate, imports and exports, disposable income, total retail sales of consumer goods, tourist visits and tourism income. The data can be obtained from the National Bureau of Statistics of China (NBSC, <http://www.stats.gov.cn/english/> (accessed on 22 August 2022)).

4. Evaluation Results and Discussion

4.1. The Evaluation Results

The evaluation mainly focuses on the adaption of different prediction models. For each model, the applicable scenes are discussed and the applicable conditions are given. By comparing the numbers of good models for each airport, the authors try to give out the reasons for failures. Considering the significance of results for the target year, four evaluation criteria with different levels of goodness are selected to assess the model suitability of different levels of goodness. Criterion 1: when the maximum mean absolute percentage error (MAPE) satisfies $\varepsilon \leq 10\%$, the model is a good model. Criterion 2: when the maximum MAPE satisfies $\varepsilon \leq 15\%$, the model is a good model. Criterion 3: when the maximum MAPE satisfies $\varepsilon \leq 20\%$, the model is a good model. Criterion 4: when the APT of the airport is more than 10 million and the maximum MAPE satisfies $\varepsilon \leq 10\%$, the model is a good model. When the APT of the airport is between 2 million and 10 million and the maximum MAPE satisfies $\varepsilon \leq 15\%$, the model is a good model. When the APT of the airport is less than 2 million and the maximum MAPE satisfies $\varepsilon \leq 20\%$, the model is a good model. The four criteria are based on different values of MAPE to set different levels of goodness which can be used to show the model suitability of different levels of goodness and to see the change in the number of predictable airports after changing the maximum allowable prediction error percentage.

4.1.1. Time Series Models

For the trend extrapolation models including linear trend, quadratic trend, cubic trend, power trend, exponential trend and logarithmic trend models, the most effective model is firstly determined for a certain airport. The coefficient of determination is used to judge the effectiveness of the models and the model with the largest coefficient of determination is selected as the model for prediction. If the coefficients of determination of all the six models are smaller than 0.8, the trend extrapolation models are deemed to be failed for prediction.

For the exponential smoothing models including quadratic exponential smoothing models and cubic exponential smoothing models, different values of the smoothing coefficients exert a tremendous influence on the predicted results. The most effective exponential smoothing model is determined by computing the MAPE of models under different smoothing coefficients. By setting the MAPE as 10%, the distribution of the number of most effective exponential smoothing models for airports under different smoothing coefficients is shown in Figure 3. For quadratic exponential smoothing models, when the smoothing coefficient is set as 0.45, the corresponding model is suitable for most airports under the given error. When the smoothing coefficient is set as 0.2, the corresponding cubic exponential smoothing model is suitable for most airports under the given error.

For the grey models, at least four points are needed to initiate the model which means that the minimum order number of a GM (1,1) is 4. By accumulating the observation sequences, the models are established and the effectiveness of the models is judged by the relative error test, correlation test and posterior difference test.

The evaluating results of the above models are shown in Table 2, from which we can see that the exponential smoothing models perform better than the other two types of models. The exponential smoothing models can be used as good models for 68% of all the 203 airports under all four evaluating criteria. The trend extrapolation models show better performances in larger airports. The grey models can predict accurately for fewer airports because only a few airports show the characteristics of exponential growth.

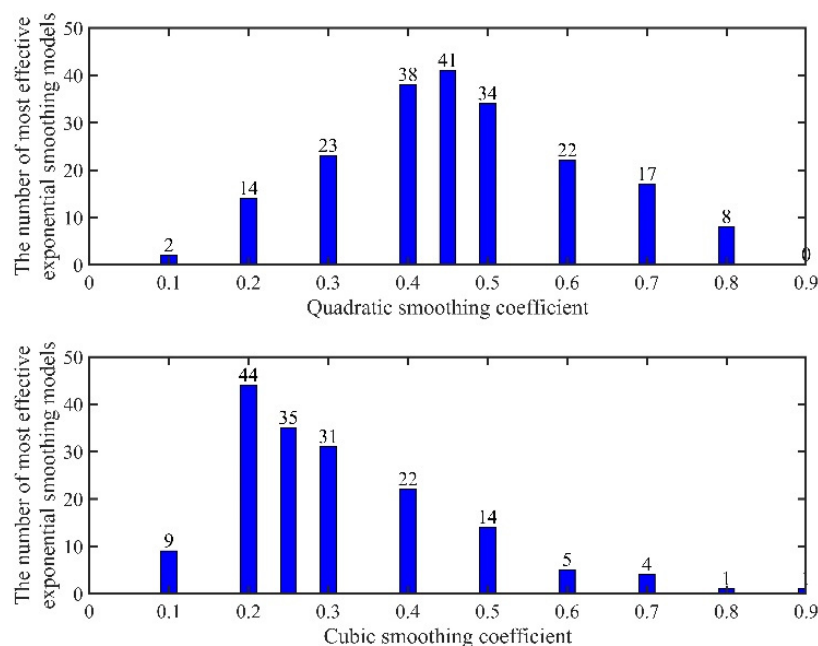


Figure 3. The distribution of the number of most effective exponential smoothing models.

Table 2. The evaluating results of the time series models.

	Trend Extrapolation Models		Exponential Smoothing Models		Grey Models	
	Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage
Criterion 1	68	33%	139	68%	15	7%
Criterion 2	89	44%	139	68%	28	14%
Criterion 3	110	54%	139	68%	41	20%
Criterion 4	95	47%	139	68%	30	15%

The applicability of the models on different levels of airports is shown in Table 3. The time series models are defined to be effective for an airport when at least one of the models is a good model under the given criterion. For the airports with APT of more than 10 million, 100% of the airports can be predicted effectively. Meanwhile, the time series models show poor applicability to airports with APT of less than 2 million. The criteria do not perform much influence on the predicted results. Considering the unavailability of the time series models, the airports with APT between 2 million and 10 million are studied in detail. There are only three airports that cannot be predicted by the time series models, which are shown in Figure 4. The Yangzhou airport (YTY) started running in 2012 and the little data make the models difficult to fit the observations well. The Yichang airport (YIH) is influenced by the Three Gorges Dam Project and its APT has shown explosive growth since 2014 when the Three Gorges Dam Project was opened to the public. The Zhangjiajie airport (DYG) develops depending much on the tourist industry. At first, the Zhangjiajie airport experienced a rapid expansion owing to the development of tourism resources. Then for a while, the Zhangjiajie airport was affected by the development of the high-speed railway. In 2017, after the visa on arrival was permitted, the APT of the airport returned to positive growth. For the airports with APT less than 2 million, 49 airports cannot be predicted by the time series models. Eighteen airports lack sufficient data. Eight airports suffer from data interruption because of expansion or relocation. The others are driven by external forces and show a lot of uncertainties like the Jiuzhaigou (JZH) airport.

Table 3. The applicability of time series models on different levels of airports.

	Criterion 1		Criterion 2		Criterion 3		Criterion 4	
	Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage
Summation of all levels	145	71%	149	73%	151	74%	150	74%
APT less than 2 million	78	60%	80	62%	81	62%	81	62%
APT between 2 million and 10 million	28	82%	30	88%	31	91%	30	88%
APT more than 10 million	39	100%	39	100%	39	100%	39	100%

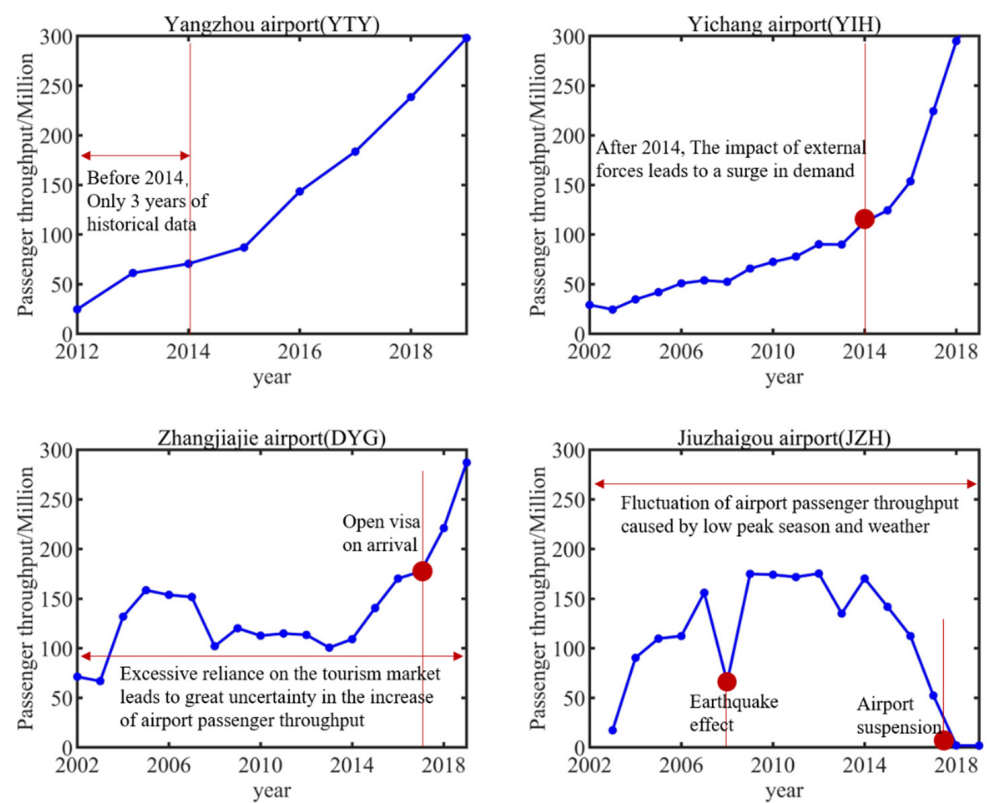


Figure 4. The four airports which are unavailable predicted by the time series models.

4.1.2. Causal Models

The accuracy of the causal models is mainly influenced by the choices of variables. The relationship between the variables and APT should be first studied. Many studies tried to solve the relationship. Wu et al. [35] analyzed the relationship between APT and indicators of socio-economic development in each province of China. The result showed that the APT was strongly positively correlated with the GDP, urbanization rate and population density and the APT per unit GDP was weakly correlated with the GDP, urbanization rate and population density. In this paper, the relationship is comprehensively researched. GDP, GDP of primary industry, GDP of secondary industry, GDP of tertiary industry, resident population, year-end population, urbanization rate, imports and exports, disposable income, total retail sales of consumer goods, tourist visits and tourism income are selected as the variables. The correlation analysis is conducted on the APT of the airport and the indicators of the province where the airport is located. Usually, when the correlation coefficient between two variables is larger than 0.8, the two variables are believed to be

highly correlated. The correlation coefficients between APT and each indicator are shown in Table 4. It is shown that the APT of an airport is highly correlated to the GDP of the city where the airport is located. In addition, the higher the magnitude of airport passenger throughput, the better the correlation between passenger throughput and macro indicators.

Table 4. The correlation coefficients between APT and indicators.

	GDP	GDP of Primary Industry	GDP of Secondary Industry	GDP of Tertiary Industry	Resident Population	Year-End Population	Urbanization Rate	Imports and Exports	Disposable Income	Total Retail Sales of Consumer Goods	Tourist Visits	Tourism Income
All the airport APT more than 10 million	74%	67%	53%	85%	48%	41%	59%	36%	85%	84%	80%	76%
APT more than 2 million	100%	72%	87%	100%	85%	79%	79%	67%	100%	100%	92%	97%
APT more than 1 million	97%	78%	79%	97%	75%	68%	73%	51%	97%	97%	93%	96%
	91%	75%	68%	93%	62%	58%	68%	47%	93%	92%	89%	89%

To exclude the influence of correlation between the indicators, the correlation between GDP and the other indicators is also studied and the results are shown in Table 5. It is shown that GDP is strongly correlated with the GDP of secondary industry, GDP of tertiary industry, disposable income and total retail sales of consumer goods.

Table 5. The correlation coefficients between GDP and the other indicators.

	GDP	GDP of Primary Industry	GDP of Secondary Industry	GDP of Tertiary Industry	Resident Population	Year-End Population	Urbanization Rate	Imports and Exports	Disposable Income	Total Retail Sales of Consumer Goods	Tourist Visits	Tourism Income
All the airport APT more than 10 million	99%	85%	90%	91%	55%	59%	57%	46%	87%	86%	76%	75%
APT more than 2 million	100%	87%	100%	100%	85%	82%	69%	69%	100%	100%	90%	95%
APT more than 1 million	100%	92%	100%	100%	78%	75%	70%	63%	100%	100%	88%	89%
	94%	88%	93%	94%	67%	69%	64%	58%	93%	94%	82%	81%

For the regression models, the effectiveness of the models is determined by the coefficient of determination. The more the coefficient of determination is close to 1, the better the model is. Usually, the model is regarded as an effective model when the coefficient of determination is larger than 0.8. In this paper, the F inspection value and t inspection value at the significance level of 0.95 are also employed.

For unary linear regression models, the results of all 12 indicators are calculated. As shown in Table 6, the predicting results of unary linear regression models do not perform well although by using the coefficient of determination the variables are highly correlated to the APT of most airports. It can be found that GDP, GDP of tertiary industry, disposable income, total retail sales of consumer goods, tourist visits and tourism income highly affect the effectiveness of all the models.

For stepwise regression models, by using all 12 indicators as the input variables, the predicting performance may be influenced by the correlation between the indicators although the significance level of the model is guaranteed by the model’s building process. Therefore, GDP, GDP of tertiary industry, disposable income, total retail sales of consumer goods, tourist visits and tourism income are selected as the limited inputs to build the models as well. As is shown in Table 7, the strategy of using limited inputs performs better.

Table 6. The evaluating results of unary linear regression models.

	GDP	GDP of Primary Industry	GDP of Secondary Industry	GDP of Tertiary Industry	Resident Population	Year-End Population	Urbanization Rate	Imports and Exports	Disposable Income	Total Retail Sales of Consumer Goods	Tourist Visits	Tourism Income
Criterion 1	17	6	4	21	3	5	0	4	27	20	13	8
Criterion 2	31	9	7	32	4	12	0	5	37	28	16	19
Criterion 3	43	17	13	45	7	15	1	9	49	39	26	26
Criterion 4	30	12	6	32	3	7	1	5	37	26	15	17
Number of airports with a coefficient of determination larger than 0.8	157	157	148	162	93	101	90	106	147	159	140	136
Number of airports whose models pass through the inspection	122	102	109	122	27	55	16	75	106	120	100	90

Table 7. The evaluating results of stepwise regression models.

	Using All 12 Indicators as the Input Variables	Using Limited Indicators as the Input Variables
Criterion 1	29	22
Criterion 2	42	33
Criterion 3	62	50
Criterion 4	47	39

For the hybrid regression causal models, each model is built using the most correlated variable and the evaluating results are shown in Table 8.

Table 8. The evaluating results of the hybrid regression causal model.

Criterion	Number of Effective Models
Criterion 1	29
Criterion 2	42
Criterion 3	62
Criterion 4	47

For the elastic coefficient models, GDP, GDP of tertiary industry, disposable income, total retail sales of consumer goods, tourist visits and tourism income are selected as the correlated factors and the evaluating results are shown in Table 9. The results indicate that the models using GDP, GDP of tertiary industry, disposable income and total retail sales of consumer goods perform better.

Table 9. The evaluating results of elastic coefficient models.

	GDP	GDP of Tertiary Industry	Disposable Income	Total Retail Sales of Consumer Goods	Tourist Visits	Tourism Income
Criterion 1	23	24	28	25	15	19
Criterion 2	35	33	41	35	22	29
Criterion 3	51	43	56	44	32	38
Criterion 4	36	34	41	34	24	29

For the proposed elastic-like scale models, the GDP, tourist visits, and resident population are chosen as the indicators. The effectiveness of the regression models is defined as any one of the unary linear regression models, stepwise regression models and hybrid regression models that can predict the APT of an airport effectively. The effectiveness of the elastic coefficient models is defined as any one of the elastic coefficient models that can predict the APT of an airport effectively. Then, the performances of regression models,

elastic coefficient models and the elastic-like scale models can be compared, as shown in Table 10.

Table 10. The comparison of the causal models.

	Regression Models		Elastic Coefficient Models		Scale Model Using GDP as the Indicator		Scale Model Using Tourist Visits as the Indicator		Scale Model Using GDP as the Indicator	
	Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage
Criterion 1	79	39%	122	60%	99	49%	74	36%	80	39%
Criterion 2	101	50%	132	65%	106	52%	85	42%	99	49%
Criterion 3	123	61%	141	69%	119	59%	95	47%	110	54%
Criterion 4	114	56%	138	68%	110	54%	89	44%	98	48%

As shown in Table 11, the applicability of the causal models on different airport levels shows that 80% of airports can use the causal models to predict. For airports with APT of more than 2 million, the models can be used to predict 100% of airports. For those that cannot be predicted, the APT of the airports shows characteristics including inconsistent trending of APT and economic development, lower correlation between the APT and economic indicators and incoherent observations because of stopping the service, as shown in Figure 5.

Table 11. The applicability of causal models on different levels of airports.

	Criterion 1		Criterion 2		Criterion 3		Criterion 4	
	Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage
Summation of all levels	146	72%	152	75%	153	75%	152	75%
APT less than 2 million	79	61%	83	64%	84	65%	84	65%
APT between 2 million and 10 million	29	85%	30	88%	30	88%	30	88%
APT more than 10 million	38	97%	39	100%	39	100%	38	97%

4.1.3. Market Share Methods

For all 203 airports, each airport is located in a city without other airports except in the cities of Beijing, Shanghai and Chengdu, which means that for most of the airports, the market shares of the airports are the same as the market shares of the cities when using the market share method. The market share methods can be used to calculate the APT of airports accounting for the province’s market share, the APT of airports accounting for the nation’s market share, as well as the APT of airports accounting for the city agglomeration’s market share. Since the 19 city agglomerations in China only cover 121 airports, the corresponding method is suitable for 121 airports. The market shares are fitted by using the time series models and the evaluating results are shown in Table 12.

By evaluating the criteria proposed, the applicability of the market share methods on different airport levels is listed in Table 13. Generally, the market share methods are effective in 70% of airports. Under different criteria, the effectiveness does not fluctuate a lot. By checking the applicability of the market share methods on different levels of airports, the market share methods can predict the airports with APT of more than 10 million accurately except for the Ningbo airport under Criterion 1. For airports with APT between 2 million and 10 million, the airports can all be predicted except for the Yangzhou airport, Yichang airport and Zhangjiajie airport under Criterion 1. Overall, 60% of the airports with APT less than 2 million can be predicted effectively. For the airports which are cannot be predicted by the market share methods, the observations of 18 airports are too less to fit the models,

and the others fluctuate disorderly because their own development regularity is not strong enough like Baotou airport (BAV) shown in Figure 6. The unpredictable airports are all located in remote areas. Comparing the performances of time series models with market share methods, the predictive abilities of the two kinds of methods are nearly the same, which is shown in Table 14. The airports which can be predicted by the time series models but cannot be predicted by the market share methods are shown in Table 15.

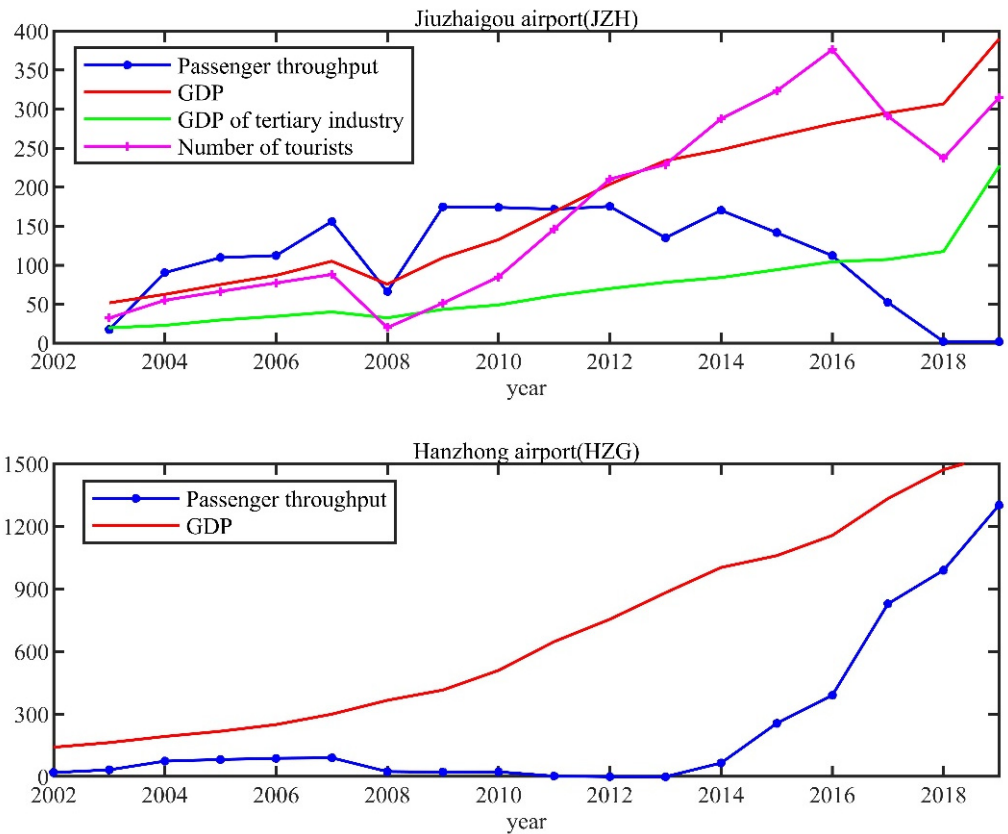


Figure 5. The airports cannot be predicted by the causal models.

Table 12. The evaluating results of market share methods.

	Province’s Market Share Method		Nation’s Market Share Method		City Agglomeration’s Market Share Method	
	Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage
Criterion 1	124	61%	128	63%	68	56%
Criterion 2	134	66%	133	66%	73	60%
Criterion 3	139	68%	135	67%	75	62%
Criterion 4	136	67%	134	66%	71	59%

4.1.4. Analogy-Based Method

According to the constructing process of the analogy-based method proposed in Section 2, the APT of all 203 airports can be used as the database for analogy. The target airports for analogy are determined by the levels of the focusing airports. The evaluation results of the analogy-based method on different levels of airports are shown in Table 16.

Table 13. The applicability of the market share method on different levels of airports.

	Criterion 1		Criterion 2		Criterion 3		Criterion 4	
	Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage
Summation of all levels	146	72%	152	75%	153	75%	152	75%
APT less than 2 million	79	61%	83	64%	84	65%	84	65%
APT between 2 million and 10 million	29	85%	30	88%	30	88%	30	88%
APT more than 10 million	38	97%	39	100%	39	100%	38	97%

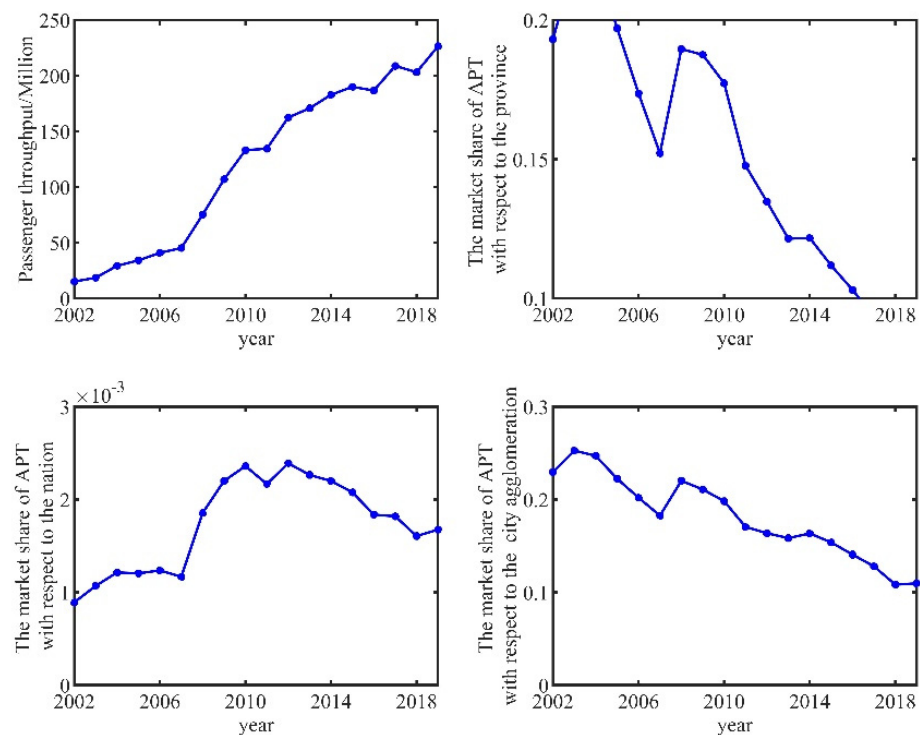


Figure 6. Market share of Baotou airport (BAV).

Table 14. Comparison of the time series models and the market share methods.

Methods	Criterion 1		Criterion 2		Criterion 3		Criterion 4	
	Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage
Time series models	145	71%	149	73%	151	74%	150	74%
Market share methods	146	72%	152	75%	153	75%	152	75%

Table 15. The airports that can be predicted by the time series models but cannot be predicted by the market share methods.

Airport	Province	Area
OHE CHG CNI	Heilongjiang Liaoning Liaoning	Northeast China
LCX JDZ	Fujian Jiangxi	East China
NZH WUA BAV	Inner Mongolia Inner Mongolia Inner Mongolia	North China
IQN JGN HZG	Gansu Gansu Shaanxi	Northwest China
PZI DIG	Szechwan Yunnan	Southwest China
KRY KJI	Xinjiang Xinjiang	Xinjiang

Table 16. The applicability of analogy-based method on different levels of airports.

	Criterion 1		Criterion 2		Criterion 3		Criterion 4	
	Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage
Summation of all levels	46	23%	67	33%	88	43%	71	35%
APT less than 2 million	18	14%	26	20%	37	28%	37	28%
APT between 2 million and 10 million	4	12%	10	29%	17	50%	10	29%
APT more than 10 million	24	62%	31	79%	34	87%	24	62%

Generally, compared with time series models, causal models and market share methods, the performance of the analogy-based method is worse, especially for the smaller airports. In addition, under different criteria, the performances are quite different, which indicates that the prediction accuracy of the analogy-based method is relatively low. By analyzing the applicable scenes of the analogy-based method, it is found that the analogy-based method is more suitable for the APT with a linear trend like Shanghai Pudong Airport (PVG) shown in Figure 7. When the increasing rate of APT is small, the predicted result will be relatively larger than the observation like Sanya Airport (SYX). When the increasing rate of APT is small, the predicted results will be relatively smaller than the observation of Nantong Airport (NTG) and Zhuhai Airport (ZUH).

4.1.5. Artificial Intelligent Models

The artificial intelligent models are built using the neural network toolbox provided by MATLAB. A feedforward neural network with 12 inputs, 1 output and 10 neurons in the hidden layer is first built. The best performance appears when the training function is set to be 'trainlm'. Compared with the other proposed models, the artificial intelligent model performs worst. Only 10% of all the airports can be predicted by the neural network. The evaluating results of the artificial intelligent model on different airport levels are shown in Table 17.

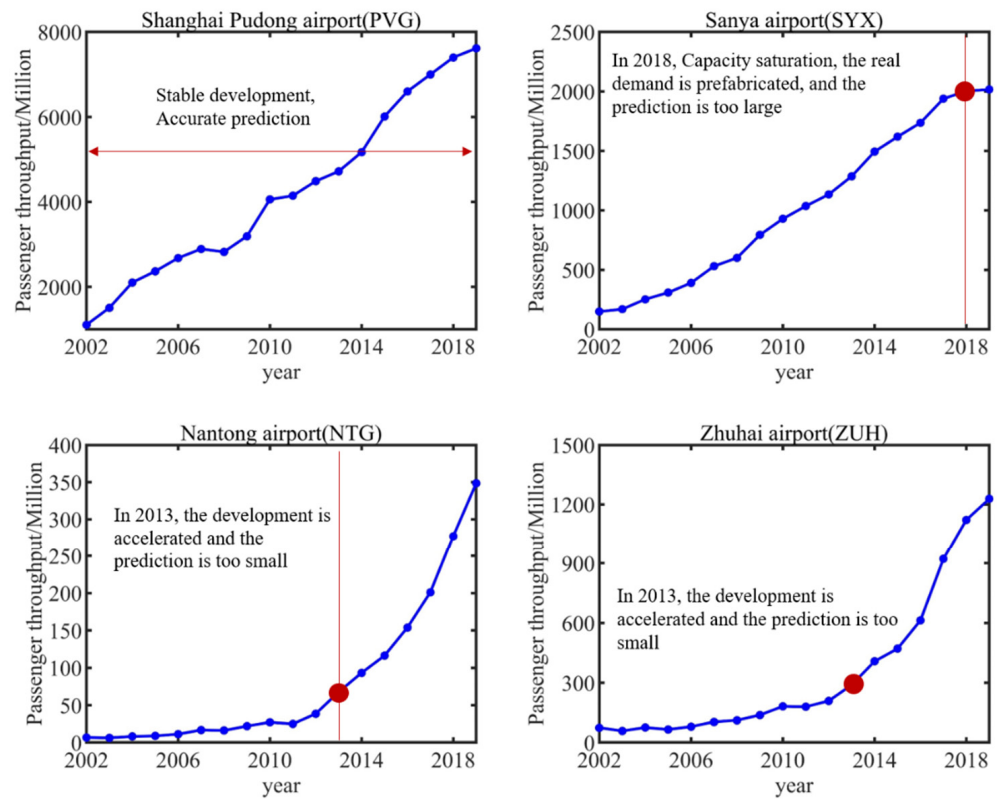


Figure 7. The applicable scenes of the analogy-based methods.

Table 17. The applicability of the artificial intelligent model on different levels of airports.

	Criterion 1		Criterion 2		Criterion 3		Criterion 4	
	Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage
Summation of all levels	16	8%	17	8%	20	10%	18	9%
APT less than 2 million	9	7%	9	7%	11	8%	11	8%
APT between 2 million and 10 million	2	6%	2	6%	2	6%	2	6%
APT more than 10 million	5	13%	6	15%	7	18%	5	13%

4.2. Analysis and Discussion of the Hybrid Method

The APT of an airport is deemed to be effectively predicted when there is at least one of five proposed models can be used to predict the APT effectively for each airport under arbitrary criteria and the hybrid five-model method is applicable. Approximately 180 of 203 tested airports can be effectively predicted and the total predictable proportion is about 88% under Criterion 1. It reveals that the hybrid method combining the time series model, the causal model, the artificial intelligent model, the market share model and the analogy-based model is applicable to most airports. The unpredictable airports of the five-model hybrid method can generally be divided into two categories. One is because there are insufficient historical data for modelling. The other one is because the APT of the airports changes abruptly owing to expansion, relocation or other kinds of external forces such as earthquakes.

For different levels of airport, all airports with APT of more than 10 million can be effectively predicted under arbitrary criteria with the five-model hybrid method. Regardless of the Zhangjiajie airport (DYG), all the airports with APT between 2 million and 10 million can be effectively predicted under Criterion 1. Under Criterion 2, all the airports with APT between 2 million and 10 million can be effectively predicted. For the airports with APT less than 2 million, 85% of the airports can be effectively predicted. By comparing the number of good models for each airport shown in Figure 8, there are usually more than three good models for the airports with APT of more than 10 million under Criterion 3. For the airports with APT between 2 million and 10 million, there is at least one good model. The applicability of all the models is promoted with the increasing APT. The applicability of all the models on different levels of airports is shown in Table 18. It reveals that large airports with high APT usually have high prediction accuracy with the hybrid evaluating method. Airports with low APT are hard to be predicted with the promoted hybrid evaluating method.

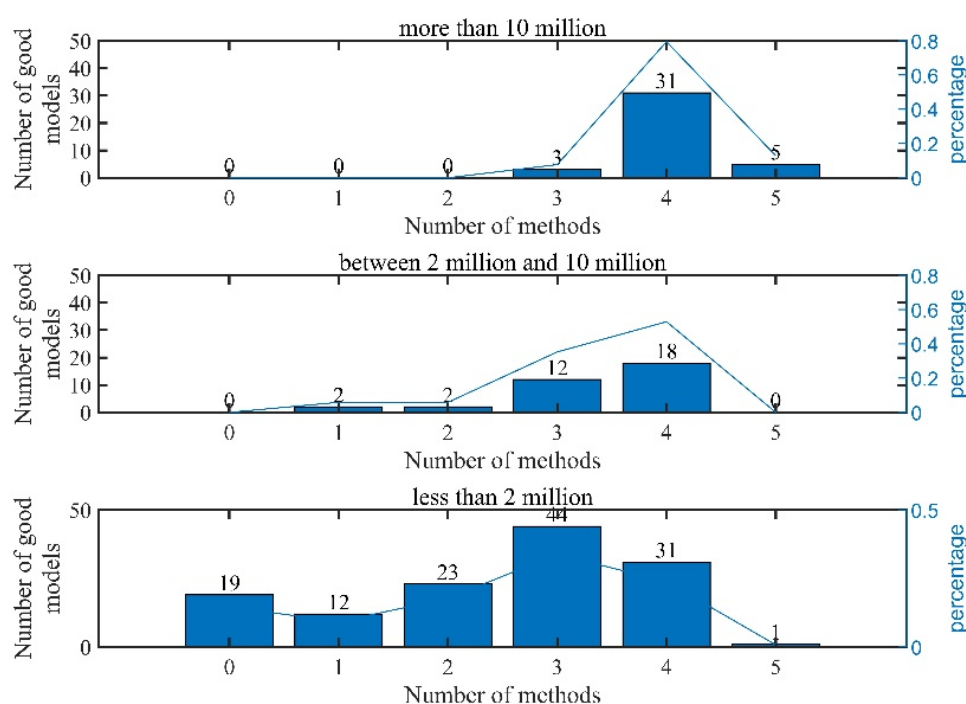


Figure 8. The distribution of the good models.

Table 18. The applicability of all the models on different levels of airports.

	Criterion 1		Criterion 2		Criterion 3		Criterion 4	
	Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage
Summation of all levels	178	88%	183	90%	184	91%	184	91%
APT less than 2 million	106	82%	110	85%	111	85%	111	85%
APT between 2 million and 10 million	33	97%	34	100%	34	100%	34	100%
APT more than 10 million	39	100%	39	100%	39	100%	39	100%

Considering the errors of the effective models, the MAPE performances of the models are given in Figure 9 and Table 19. For airports with APT of more than 10 million, the

MAPE is usually smaller than that for the small airports. The MAPEs of time series models, causal models and market share method are less than 10%.

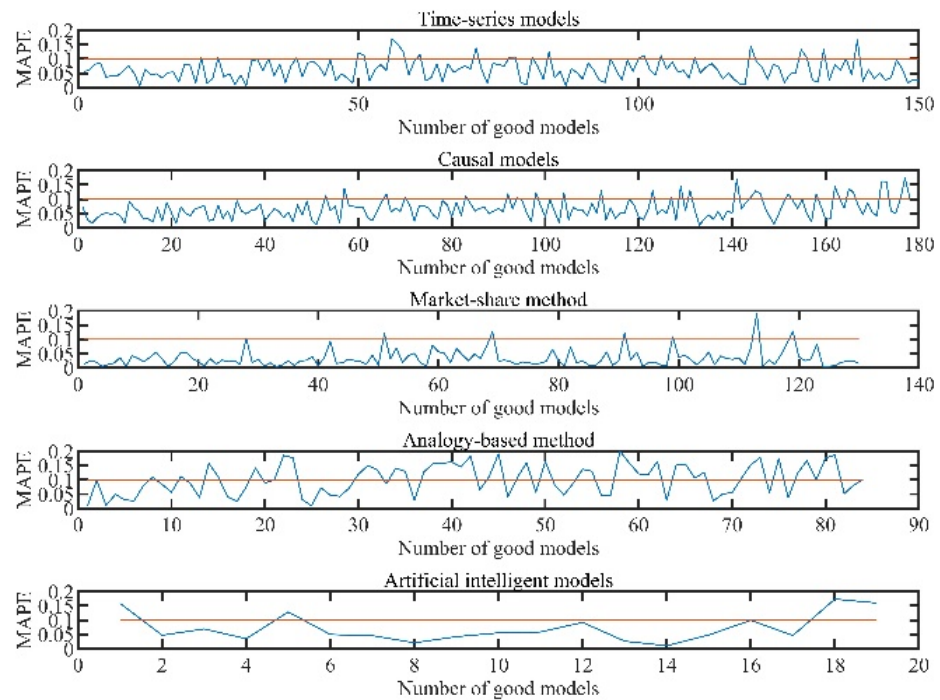


Figure 9. The MAPE distribution of the effective models.

Table 19. The prediction performances of the effective models according to MAPE.

		0–5%		5–10%		10–15%		15–20%	
		Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage
Time series models	Summation of all levels APT less than 2 million	63	31%	63	31%	22	11%	3	1%
	APT between 2 million and 10 million	19	49%	16	41%	4	10%	0	0%
	APT more than 10 million	12	35%	13	38%	4	12%	2	6%
		34	26%	34	26%	14	11%	1	1%
Causal models	Summation of all levels APT less than 2 million	61	30%	89	44%	25	12%	4	2%
	APT between 2 million and 10 million	18	46%	21	54%	0	0%	0	0%
	APT more than 10 million	12	35%	18	53%	4	12%	0	0%
		31	24%	50	38%	21	16%	4	3%

Table 19. Cont.

		0–5%		5–10%		10–15%		15–20%	
		Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage	Number of Good Models	Percentage
Market share methods	Summation of all levels	127	63%	19	9%	6	3%	1	0%
	APT less than 2 million	35	90%	3	8%	1	3%	0	0%
	APT between 2 million and 10 million	27	79%	2	6%	1	3%	0	0%
	APT more than 10 million	65	50%	14	11%	4	3%	1	1%
Analogy based methods	Summation of all levels	24	12%	22	11%	21	10%	21	10%
	APT less than 2 million	14	36%	10	26%	7	18%	3	8%
	APT between 2 million and 10 million	2	6%	2	6%	6	18%	7	21%
	APT more than 10 million	8	6%	10	8%	8	6%	11	8%
Artificial intelligent models	Summation of all levels	11	5%	5	2%	1	0%	3	1%
	APT less than 2 million	4	10%	1	3%	1	3%	1	3%
	APT between 2 million and 10 million	2	6%	0	0%	0	0%	0	0%
	APT more than 10 million	5	4%	4	3%	0	0%	2	2%

5. Conclusions

Time series models, causal models, artificial intelligent models, market share methods and analogy-based methods are all commonly used methods for APT prediction. In the paper, a hybrid method of the aforementioned methods is developed and investigated. By conducting the evaluation of five kinds of models to predict the APT of two hundred three airports in China, it is found that for most prediction scenes, the models are applicative to the short-term prediction of APT and the accuracy does not improve with the complexities of the models. Overall, 88% of the studied airports can be effectively predicted by using the evaluated prediction methods, and the MAPE is mostly within 10%. In addition, the constructed evaluation method can effectively predict the airport passenger throughput of more than 10 million. The higher the airport passenger throughput level is, the more effective the segmentation prediction methods can be. The performance of the mentioned models in this paper is bad when there are insufficient historical data for modelling, or the APT of the airports changes abruptly owing to expansion, relocation or other kinds of external forces like earthquakes. When the airport enters the stable development period, the number of available prediction models in the hybrid method is significantly increased. The results show that there is no relationship between the prediction accuracy and the complexity of prediction models. Among the five used models, the time series model, causal models and market share method usually have higher applicability than the other two models. The findings provide potential support for the selection of prediction models for the APT prediction. The prediction and evaluation methods in this paper are usually applicable to airports with high APT, stable development periods and sufficient historical

data. Further studies should be made to build prediction and evaluation methods when the external environment is unstable under the influence of COVID-19 and other external forces such as earthquakes or without sufficient historical data.

Author Contributions: Conceptualization, B.C. and J.W.; methodology, B.C.; software, B.C.; validation, B.C., J.W. and X.Z.; formal analysis, X.Z.; investigation, B.C.; resources, B.C.; data curation, B.C.; writing—original draft preparation, B.C.; writing—review and editing, X.Z.; visualization, X.Z.; supervision, J.W. and X.Z.; project administration, J.W. and X.Z.; funding acquisition, J.W. and X.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the project “Civil Aviation Safety Capacity Enhancement Project”, which is supported by the Civil Aviation Administration of China.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors gratefully acknowledge the financial support provided by the project “Civil Aviation Safety Capacity Enhancement Project”, which is supported by the Civil Aviation Administration of China.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Solvoll, G.; Mathisen, T.A.; Welde, M. Forecasting air traffic demand for major infrastructure changes. *Res. Transp. Econ.* **2020**, *82*, 100873. [[CrossRef](#)]
2. Kurdel, P.; Sedláčková, A.N.; Novák, A. Analysis of Using Time Series Method for Prediction of Number of Passengers at the Airport. *J. KONBiN* **2020**, *50*, 203–216. [[CrossRef](#)]
3. Cheng, L.; Xiao, M. A Review of Research on Airline Passenger Volume Forecasting. In Proceedings of the 2017 4th International Conference on Machinery, Materials and Computer (MACMC 2017), Xi’an, China, 27–29 November 2017.
4. Zhang, X. Research on forecasting method of aviation traffic based on social and economic indicators. *IOP Conf. Ser. Mater. Sci. Eng.* **2020**, *780*, 062038. [[CrossRef](#)]
5. Peng, D.; Zhang, M.; Xiao, Y.; Wang, Y. Research on Passenger Throughput Forecast of Civil Aviation Airport Based on Multi-source Data. *J. Phys. Conf. Ser.* **2022**, *2179*, 012027. [[CrossRef](#)]
6. Bastola, D.P. Air Passenger Demand Model (APDM): Econometric Model For Forecasting Demand In Passenger Air Transports In Nepal. *Int. J. Acad. Res. Psychol.* **2017**, *1*, 238–242.
7. Jin, F.; Li, Y.; Sun, S.; Li, H. Forecasting air passenger demand with a new hybrid ensemble approach. *J. Air Transp. Manag.* **2019**, *83*, 101744. [[CrossRef](#)]
8. Xie, G.; Wang, S.; Lai, K.K. Short-term forecasting of air passenger by using hybrid seasonal decomposition and least squares support vector regression approaches. *J. Air Transp. Manag.* **2014**, *37*, 20–26. [[CrossRef](#)]
9. Xu, S.; Chan, H.K.; Zhang, T. Forecasting the demand of the aviation industry using hybrid time series SARIMA-SVR approach. *Transp. Res. Part E Logist. Transp. Rev.* **2018**, *122*, 169–180. [[CrossRef](#)]
10. Gunter, U.; Zekan, B. Forecasting air passenger numbers with a GVAR model. *Ann. Tour. Res.* **2021**, *89*, 103252. [[CrossRef](#)]
11. Scarpel, R.A. Forecasting air passengers at São Paulo International Airport using a mixture of local experts model. *J. Air Transp. Manag.* **2013**, *26*, 35–39. [[CrossRef](#)]
12. Zhou, W.; Cheng, Y.; Ding, S.; Chen, L.; Li, R. A grey seasonal least square support vector regression model for time series forecasting. *ISA Trans.* **2020**, *114*, 82–98. [[CrossRef](#)]
13. Wong, H.L. Time Series Forecasting with Stochastic Markov Models Based on Fuzzy Set and Grey Theory. *Appl. Mech. Mater.* **2015**, *764–765*, 975–978. [[CrossRef](#)]
14. Liang, X.; Zhang, Q.; Hong, C.; Niu, W.; Yang, M. Do Internet Search Data Help Forecast Air Passenger Demand? Evidence From China’s Airports. *Front. Psychol.* **2022**, *13*, 809954. [[CrossRef](#)]
15. Li, X.; Groot, M.D.; Bck, T. Using forecasting to evaluate the impact of COVID-19 on passenger air transport demand. *Decis. Sci.* **2021**. [[CrossRef](#)]
16. Barczak, A.; Dembińska, I.; Rozmus, D.; Szopik-Depczyńska, K. The Impact of COVID-19 Pandemic on Air Transport Passenger Markets—Implications for Selected EU Airports Based on Time Series Models Analysis. *Sustainability* **2022**, *14*, 4345. [[CrossRef](#)]
17. Karlaftis, M.G.; Zografos, K.G.; Papastavrou, J.D.; Charnes, J.M. Methodological Framework for Air-Travel Demand Forecasting. *J. Transp. Eng.* **1996**, *122*, 96–104. [[CrossRef](#)]
18. Maldonado, J. Strategic Planning—An Approach to Improving Airport Planning under Uncertainty. Master’s Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 1990.

19. Mierzejewski, E.A. *A New Strategic Urban Transportation Planning Process*; Center for Urban Transportation Research, University of South Florida: Tampa, FL, USA, 1995.
20. Odeck, J.; Welde, M. The accuracy of toll road traffic forecasts: An econometric evaluation. *Transp. Res. Part A Policy Pract.* **2017**, *101*, 73–85. [[CrossRef](#)]
21. Flyvbjerg, B.; Holm, M.K.S.; Buhl, S.L. How (In)accurate Are Demand Forecasts in Public Works Projects?: The Case of Transportation. *J. Am. Plan. Assoc.* **2005**, *71*, 131–146. [[CrossRef](#)]
22. Nicolaisen, M.S.; Driscoll, P.A. *Ex-Post Evaluations of Demand Forecast Accuracy: A Literature Review*. *Transp. Rev.* **2014**, *34*, 540–557. [[CrossRef](#)]
23. Djakaria, I. Djalaluddin Gorontalo Airport Passenger Data Forecasting with Holt's-Winters' Exponential Smoothing Multiplicative Event-Based Method. *J. Phys. Conf. Ser.* **2019**, *1320*, 012051. [[CrossRef](#)]
24. Kochkina, E.; Radkovskaya, E.; Denezhkina, K. Analysis and forecasting of performance indicators of air transport facilities. *E3S Web Conf.* **2021**, *291*, 08009. [[CrossRef](#)]
25. Priyadarshana, M. Modeling Air Passenger Demand in Bandaranaike International Airport, Sri Lanka. *J. Bus. Econ. Policy* **2015**, *2*, 146–151.
26. Zhang, F.; Graham, D.J. Air transport and economic growth: A review of the impact mechanism and causal relationships. *Transp. Rev.* **2020**, *40*, 506–528. [[CrossRef](#)]
27. Goldfarb, R.S.; Stekler, H.O.; David, J. Methodological issues in forecasting: Insights from the egregious business forecast errors of late 1930. *J. Econ. Methodol.* **2005**, *12*, 517–542. [[CrossRef](#)]
28. Wang, J.Z.; Wang, J.J.; Zhang, Z.G.; Guo, S.P. Forecasting stock indices with back propagation neural network. *Expert Syst. Appl.* **2011**, *38*, 14346–14355. [[CrossRef](#)]
29. Jiang, X.; Zhang, Y.; Li, Y.; Zhang, B. Forecast and analysis of aircraft passenger satisfaction based on RF-RFE-LR model. *Sci. Rep.* **2022**, *12*, 11174. [[CrossRef](#)]
30. Cheng, L.; Liu, L.; Xiao, B. Artificial Neural Networks Method for Predicting the Airport Passenger Throughput. *Aeronaut. Comput. Tech.* **2000**, *30*, 8–11.
31. Wang, Z.; Song, W.K. Sustainable airport development with performance evaluation forecasts: A case study of 12 Asian airports. *J. Air Transp. Manag.* **2020**, *89*, 101925. [[CrossRef](#)]
32. Profillidis, V. An ex-post assessment of a passenger demand forecast of an airport. *J. Air Transp. Manag.* **2012**, *25*, 47–49. [[CrossRef](#)]
33. Liu, S.Y.; Liu, S.; Tian, Y.; Sun, Q.L.; Tang, Y.Y. Research on Forecast of Rail Traffic Flow Based on ARIMA Model. *J. Phys. Conf. Ser.* **2021**, *1792*, 012065. [[CrossRef](#)]
34. Kitsou, S.P.; Koutsoukis, N.S.; Chountalas, P.; Rachaniotis, N.P. International Passenger Traffic at the Hellenic Airports: Impact of the COVID-19 Pandemic and Mid-Term Forecasting. *Aerospace* **2022**, *9*, 143. [[CrossRef](#)]
35. Xiangli, W.U.; Man, S. Air transportation in China: Temporal and spatial evolution and development forecasts. *J. Geogr. Sci.* **2018**, *28*, 1485–1499.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.