Forecasting Post-Epidemic Air Passenger Flow Among Hub Cities in China Based on PLS and GA-SVR Model

Guo-Dong Li, Civil Aviation University of China, China

Wen-Shan Liu, Civil Aviation University of China, China
iD https://orcid.org/0009-0000-8851-4708

Sang-Bing Tsai, International Engineering and Technology Institute, Hong Kong*
iD https://orcid.org/0000-0001-6988-5829

ABSTRACT

For stakeholders and companies involved in civil aviation to plan and make wise decisions, accurate estimates of air passenger flow are necessary. The conclusions drawn from this model serve as an invaluable resource for pertinent choices. The top hub cities in mainland China's air travel network are predicted using a variable weight combination model that combines PLS and GA-SVR. According to the test results, the model developed in this study improves prediction accuracy. This demonstrates how well detailed information about social and economic development can be gleaned from linear development patterns and nonlinear fluctuation rules. Predictions can be made with more accuracy and a better fit as a result. Over the next five years, it is predicted that over 300 million passengers will fly between the top hub cities in mainland China, an increase of 6.51% per year. The growth in passenger traffic varies significantly between different routes. The routes from Beijing to Shanghai and Shanghai to Shenzhen saw the most travellers, while the Beijing-Chengdu route saw the fastest growth in traveller numbers. The study's findings provide useful advice for civil aviation businesses and people involved in their decision-making, fostering growth in the sector during the post-pandemic period.

KEYWORDS

Air Passenger Transportation Market Demand, Genetic Algorithm, Partial Least Square Regression, Simulation Method, Support Vector Regression

1. INTRODUCTION

1.1. Research Background and Significance

Global civil aviation has seen substantial expansion in recent years as a result of improvements in economic globalization and air transportation infrastructure. Since 1978, when China started its economic reform and opening up, the country's civil aircraft transportation has had a brisk yearly

*Corresponding Author

increase of 16.3%. According to L. Zhang et al. (2021), this growth rate is much larger than that of other modes of transportation. However, the unexpected COVID-19 epidemic in late 2019 had a significant global social and economic impact. Traffic restrictions have been in place all over the world for an exceptionally long time because of COVID-19, an illness that is spread by inhaling the virus. These limitations set off a series of events that affected many areas, including the global supply chain, and led to a severe economic downturn. The worldwide aviation industry as well as China suffered from the 3.4% decline in the global economy that occurred in 2020.

The COVID-19 crisis has had an effect on the aviation sector and has had a substantial impact on the passenger market for civil aviation. This has led to two significant obstacles for industry's growth in the post-epidemic age. Managing the complicated and uncertain business environment of civil aviation firms has gotten harder, according to Pereira et al. (2021). Second, because of sudden outside shocks, the industry is currently subject to severe regulations. According to Shahul Hameed et al. (2022), the increasing complexity will demand more flexible decision-making from companies. After the pandemic, L. Zhang et al. (2021) discovered sizable changes in the market for civil aviation passenger travel. In China, the scale and popularity of passenger air travel have altered, resulting in greater disparities across various times and areas and a much higher level of unpredictability.

Airlines and airports must be more prepared to handle the complex passenger market and post-epidemic commercial climate if they are to successfully navigate it. Wierzbinski et al. (2023) states that precise market data collection and analysis are crucial for managing enterprise knowledge. This is a crucial element that aids companies in being more competitive and succeeding in the post-pandemic market rivalry. In order to make adaptable strategic decisions and safeguard themselves from increasing business complexity, airlines and airports need be aware of their whole target market and each segment's capability. Additionally, those who work in the civil aviation sector (such as government representatives, regulators, financiers, etc.) need to be fully aware of the industry's projected development rate and size. With the use of this information, they may assess the value of their assets, come up with investment plans, get funding, determine their financial capability, and choose whether or not to apply for government assistance. The implementation of the "New Ten Principles" and other policies by the end of 2022 is anticipated to help the Chinese civil aviation sector enter a phase of recovery and high-quality development in 2023. To support and develop China's air passenger transport business, a strong air transportation network connecting key cities is essential. For Chinese aviation companies and stakeholders to support the recovery and development of the industry, accurate forecasting of the volume and growth of air travel is essential.

The ambiguity and instability of air travel make it difficult to predict the demand for civil aviation passenger transport after an epidemic. As a result, we are faced with uncertainty. According to research by Kumbure et al. (2022) and Wu and Xiong (2021), earlier forecast techniques that depended on traditional statistical models and a single model are unable to accurately reflect the current situation of the civil aviation industry. It is crucial to rely on models that are more adapted for complicated and quickly changing scenarios in order to perform future study.

1.2. Literature Review

Time series models, regression models, and machine learning techniques have all been used in previous studies on the forecasting of the air transportation market. However, with little study being done at the route level, the main emphasis of these studies has been on predicting the national and airport levels. Studies that examine the relationship between the expansion of the economy and the volume of air travel, however, tend to concentrate on only one aspect of the economy as the driving force behind changes in the volume of air travel.

Several researchers have examined the general trajectory of national demand in the passenger and cargo transportation market on a national scale. For instance, Chen and Li (2022) employed the extreme learning machine and particle swarm optimization algorithm to forecast the volume of passenger traffic in China's civil aviation sector. In their study, Y. Zhang et al. (2021) employed a

comprehensive model that integrated various forecasting techniques, including the gray time series model, gray relational system model, exponential smoothing model, autoregressive moving average model, multiple linear regression model, and partial least square model. The objective was to predict the demand for China's air transportation market, specifically focusing on aircraft demand. Shen et al. (2019) employed system dynamics methodology to forecast the aggregate demand of the Chinese air transportation market, considering both internal and external factors. Furthermore, certain scholars have directed their attention towards analyzing the trajectory of passenger and cargo throughput within airports or groups of airports. Liang et al. (2017) employed the fuzzy C-means algorithm to forecast the airport passenger throughput after conducting seasonal adjustment and noise decomposition of the data. Moreover, existing studies on air flow prediction at the route level primarily focus on investigating the factors that influence these flows. For instance, Li et al. (2020) employed a gravity model to forecast the market demand for air passenger transportation.

All the studies above describe a single economic scale as the background for changes in air travel passenger volume and report only a single total amount. Nevertheless, the COVID-19 pandemic has impacted not only on economic growth but also had significant chain reactions across the supply chain, social consumption, population, and other fields. Meanwhile, studies by Hanson et al. (2022) and Truong (2021) demonstrate that COVID-19's impact on air transportation is influenced by socio-economic factors such as income and population status. Economic and social trends also affect air passenger flow, making it necessary to consider comprehensive development of the social economy in predicting post-epidemic air passenger flow. And should continue with a thorough description of trends and fluctuations.

This study reevaluates the relationship between air travel passenger numbers and general economic and social development, in contrast to prior literature, which concentrated on the national or airport economic scale. This study intends to produce more accurate forecasting results by providing a more thorough description of the socioeconomic dynamics surrounding changing air travel passenger counts. This study provides a straightforward method for predicting air passenger flow based on the economic and social development of the cities at both ends of the route. Applying it to different routes in mainland China allows us to eventually obtain insight into how the air transportation market is predicted to evolve among China's hub cities.

1.3. Research Methods

The general expansion of the social economy exhibits a linear trend with non-linear deviations. It is difficult to precisely discern the economic and social development law that both sets of information are trying to convey, nevertheless, when studying it holistically. Therefore, in order to extract, fit, and predict the two portions of the data, we must use a combination forecasting model. Due to its capacity to extract a wider range of original sequence information and improve prediction accuracy, the combined prediction model is used increasingly frequently. Bates and Granger made the original argument for the combined prediction model in 1969. By using a linear or nonlinear combination of the aforementioned models, the model efficiently makes use of the independent information that is contained in each individual model. By using posteriori knowledge to make adaptive adjustments, the variable weight combination model has the potential to improve its prediction accuracy in comparison to the fixed weight mixed model. Due to the variable weight combination model's ongoing optimization with the lengthening original sequence, it has become more widely used. Notable studies by Zhang (2018) and Zhang et al. (2022) have contributed to its popularity. Certain researchers in China have undertaken investigations into the application of a fixed weight combined model for the purpose of forecasting air passenger flows. Notable examples include the studies conducted by Y. Zhang et al. (2021) and Liang et al. (2017). According to Zhang (2018), the variable weight combined model is deemed more appropriate for predicting time series data due to the presence of both linear trend and volatility in the development of air passenger flows. Therefore, this study utilizes a flexible

weight combination prediction model to fully grasp the link between the overall growth of social and economic development and air travel.

It is required to employ a linear regression model to measure the association between air passenger flow and the linear trend of economic and social growth. Most of the variables have similar development trends since the measurement of overall socio-economic development needs to include a number of linked elements, such as economic scale, population size, social consumption, and others. This study chooses the partial least squares regression model (hereinafter referred to as PLS) to extract the main characteristics of economic and social development, including the development of air transportation, in order to avoid the impact of multicollinearity on the accuracy of the model, and finally more accurately measure the relationship between air transportation passenger flow and economic and social development.

It is required to utilize a machine learning model that is better suited for complicated scenarios in order to extract the association between air passenger flow and economic and social volatility. Using tiny data sets, this study continues the projection for the annual air passenger flow. It is appropriate to use the support vector regression model (hereinafter referred to as SVR) in this study to describe the effect of the fluctuation law of social and economic development on air passenger flow because it is suitable for small data sets, has strong generalization ability, stability, and high prediction accuracy. This study employs a genetic algorithm (hence referred to as GA) to optimize the parameters for building an SVR model in order to get the best performance of fitting the present data. The best SVR model parameters are obtained as a result.

This study will create multiple city pairs of the four major hub cities in the multi-airport system of mainland China based on the construction of the variable weight combination forecasting model composed of PLS and GA-SVR, and use the model to predict the size of the air passenger transport market between the two cities in these cities. The analysis and discussion of the prediction outcomes is done towards the end. The hub cities listed above include Chengdu, Beijing, Shanghai, Guangzhou, Shenzhen, and Chongqing.

This study used a simulation method to develop a unique and efficient air passenger flow model. This research aims to inform government, airport, and aviation construction and operational strategies.

2. MODEL CONSTRUCTION

2.1. Model Building Method

This section describes the specific steps for building a model in this study. The flow chart of the model building method is shown in Fig. 1. The combination prediction model's implementation procedures are as follows:

(1) The data must be preprocessed, divided, and standardized.
(2) Incorporate training data into PLS model, extract principal components for regression, and obtain fitting model, fitting outcomes, and test results.
(3) Use cross-validation and GA to obtain the best values for the penalty, tolerance, and kernel parameters after incorporating the training data into the SVR model.
(4) Incorporate the training data into the SVR model to obtain test and fitting results.
(5) Calculate the single model's parameters for each time point, incorporate the single model's results into the variable weight combined model along with them, obtain the variable weight combined prediction model's fitting and test results, and then restore the results in accordance with the initial standardized path.
(6) Determine the evaluation index value of each model's predictive impact and conduct a comparative study.

Figure 1. Model construction flow chart
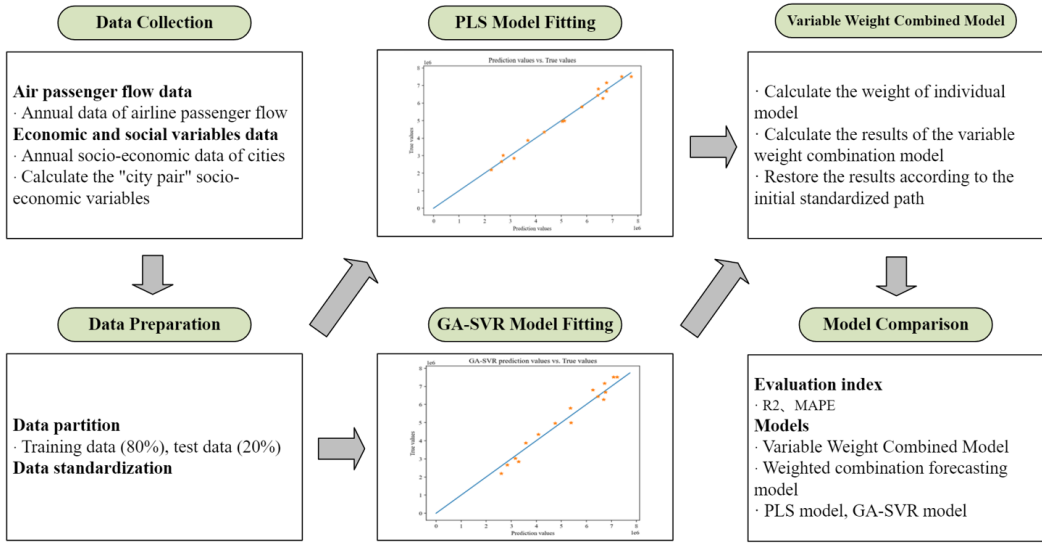
2.2. Variable Weight Combined Model and Weight Calculation

The utilization of the variable weight combined model is employed to address real-world issues. The model can be mathematically represented as follows:

$$\hat{y}_t = \sum_{i=1}^{n} w_{it} \hat{y}_{it} \, (t = 1, 2, \cdots, m)$$
$$\sum_{i=1}^{n} w_{it} = 1 \tag{1}$$

Let *n* represent the quantity of individual models and *m* denote the quantity of moments. The variable $\hat{y}_t$ denotes the anticipated value of the integrated model at a given time *t*. The weight coefficient, denoted as w_{it}, signifies the importance assigned to the individual prediction model *i* at time *t*. Meanwhile, the predicted value of the individual prediction model *i* at time *t* is represented by $\hat{y}_t$. The total weight coefficients of each individual model at each time interval equals 1.

The weight calculation method employed in this study is the reciprocal variance method. The weight in the reciprocal method of variance is determined by the sum of squares of prediction errors. The sum of squares of prediction errors is the sum of squares between the predicted values and the actual values provided by the fitting model. The accuracy of a model can be assessed by calculating the sum of squares of prediction errors. A higher value indicates lower accuracy. The weight coefficients of each individual model at each moment are determined through the following calculation:

$$w_{it} = \frac{Q_{it}^{-1}}{\sum_{i=1}^{n} Q_{it}^{-1}} \, (t = 1, 2, \cdots, m)$$
$$Q_{it} = (y_t - \hat{y}_{it})^2 \tag{2}$$

In this context, y_t represents the true value of time t, while Q_{it} represents the sum of squared errors between the prediction result of a single prediction model i and the true value at time t.

Due to the inherent uncertainty in predicting future values, it is not possible to calculate the deviation between the predicted value and the actual value. To accomplish the objective of prediction, it is imperative to forecast the future weight based on the weight of each time period computed from the available data. If the weight derived from the available data does not exhibit a discernible time series pattern, the rolling calculation can be performed using the mean value of the existing weight Deng et al. (2022), this refers to the combination of historical data and predicted data to predict the future data. The calculation rule can be described as follows:

$$w_{it} = \frac{1}{m} \sum_{t=l}^{m+l-1} w_{it} (t = m+1, m+2, \cdots, m+l) \tag{3}$$

2.3. Single Model Selection

2.3.1. PLS Model

The PLS model is an extension of the principal component regression model. The problem of multiple collinearities among independent variables in multiple regression analysis can be addressed by employing the principal component regression model. This model utilizes dimension reduction techniques to extract the principal components of the independent variables. Given the aforementioned circumstances, the PLS model considers the interdependence between the primary factors and the outcome variables. This characteristic enables the PLS model to effectively tackle the challenge of multiple regression in scenarios where the sample size is limited, there is a higher number of variables, and a substantial multicollinearity exists among all the variables. The sample size is limited as it was determined based on the annual sample selected for the purpose of conducting a long-term forecast in this study. A wide range of explanatory variables are selected, with a predominant focus on macroeconomic factors, resulting in comparable patterns of development. Based on the preceding information, it is expected that the presence of multiple collinearities among variables poses a concern. Consequently, the PLS model can be employed to delineate the trajectory of demand evolution in the realm of air passenger transportation. The realization method of the PLS model is as follows:

(1) Following the process of standardizing the dependent variable y_i (where $i = 1, 2,..., k$) and the independent variable x_j (where $j = 1, 2,..., l$), the resulting dependent variable $\mathbf{Y}$ matrix and independent variable $\mathbf{X}$ matrix are constructed.

(2) The initial principal component axis vectors extracted from $\mathbf{X}$ and $\mathbf{Y}$ are denoted as $\mathbf{p}_1$ and $\mathbf{q}_1$, respectively. The resulting first principal components generated by these vectors are denoted as $\mathbf{u}_1$ and $\mathbf{v}_1$, where $\mathbf{u}_1 = \mathbf{X} \times \mathbf{p}_1$ and $\mathbf{v}_1 = \mathbf{Y} \times \mathbf{q}_1$, respectively. The optimization model is formulated in the following manner with the objective of maximizing the correlation between $\mathbf{u}_1$ and $\mathbf{v}_1$, as well as $\mathbf{u}_1$ and $\mathbf{v}_1$. This is done to ensure that $\mathbf{u}_1$ and $\mathbf{v}_1$ capture the maximum amount of variation information from the original sequence.

$$\begin{aligned} &\max \quad Cov(\mathbf{u}_1, \mathbf{v}_1) \\ &s.t. \left\| \mathbf{p}_1 \right\| = 1, \left\| \mathbf{q}_1 \right\| = 1 \end{aligned} \tag{4}$$

(3) In the mathematical optimization problem, the Lagrange multiplier method is a method to find the extreme value of the multivariate function whose variables are restricted by one or more

conditions. This method introduces a new scalar unknown number, namely Lagrange multiplier, which is the coefficient of each vector in the linear combination of the gradient of the constraint equation. By introducing the Lagrange multiplier and constructing the Lagrange function, it is possible to derive the result that $\mathbf{p}_1$ represents the eigenvector associated with the maximum eigenvalue of $\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X}$, while $\mathbf{q}_1$ represents the eigenvector associated with the maximum eigenvalue of $\mathbf{Y}^T\mathbf{X}\mathbf{X}^T\mathbf{Y}$. Through the implementation of supplementary computations, it is possible to ascertain the initial principal component pair, denoted as $\mathbf{u}_1$ and $\mathbf{v}_1$. The regression equations for $\mathbf{X}$ and $\mathbf{Y}$ in relation to $\mathbf{u}_1$ and $\mathbf{v}_1$ can be expressed as follows: $\mathbf{X} = \mathbf{u}_1\mathbf{c}_1^T + \mathbf{X}_0$ and $\mathbf{Y} = \mathbf{v}_1\mathbf{d}_1^T + \mathbf{Y}_0'$. Within this set of variables, $\mathbf{X}_0$ and $\mathbf{Y}_0'$ can be identified as residual matrices, while $\mathbf{c}_1^T$ and $\mathbf{d}_1^T$ represent the regression coefficient matrices associated with the primary components $\mathbf{u}_1$ and $\mathbf{v}_1$, respectively. Given the correlation between $\mathbf{u}_1$ and $\mathbf{v}_1$, it becomes feasible to establish the association between $\mathbf{X}$ and $\mathbf{Y}$ by examining the association between $\mathbf{u}_1$ and $\mathbf{Y}$. This can be expressed as follows: $\mathbf{Y} = \mathbf{u}_1\mathbf{e}_1^T + \mathbf{Y}_0$.

(4) The outcomes of the least squares method are as follows:

$$\mathbf{c}_1 = \frac{\mathbf{X}^T\mathbf{u}_1}{\left\|\mathbf{u}_1\right\|^2}, \mathbf{d}_1 = \frac{\mathbf{Y}^T\mathbf{v}_1}{\left\|\mathbf{v}_1\right\|^2}, \mathbf{e}_1 = \frac{\mathbf{Y}^T\mathbf{u}_1}{\left\|\mathbf{u}_1\right\|^2} \tag{5}$$

The relationship between $\mathbf{p}_1$ and $\mathbf{c}_1$ can also be determined as follows:

$$\mathbf{p}_1^T\mathbf{c}_1 = \mathbf{p}_1^T \frac{\mathbf{X}^T\mathbf{u}_1}{\left\|\mathbf{u}_1\right\|^2} = \frac{\mathbf{u}^T\mathbf{u}_1}{\left\|\mathbf{u}_1\right\|^2} = 1 \tag{6}$$

(5) Re-evaluate the values of $\mathbf{X}$ and $\mathbf{Y}$ by following steps (1) through (4), utilizing the components of $\mathbf{X}$ and $\mathbf{Y}$ that are not accounted for by the principal component $\mathbf{u}_1$ (referred to as $\mathbf{X}_0$ and $\mathbf{Y}_0$) as the updated $\mathbf{X}$ and $\mathbf{Y}$. The variable r represents the cumulative count of primary components that have been extracted. The initial values of $\mathbf{X}$ and $\mathbf{Y}$ can be expressed as follows, considering the aforementioned calculations.

$$\mathbf{X} = \mathbf{u}_1\mathbf{c}_1^T + \mathbf{u}_2\mathbf{c}_2^T + \cdots + \mathbf{u}_r\mathbf{c}_r^T + \mathbf{X}_0$$
$$\mathbf{Y} = \mathbf{u}_1\mathbf{e}_1^T + \mathbf{u}_2\mathbf{e}_2^T + \cdots + \mathbf{u}_r\mathbf{e}_r^T + \mathbf{Y}_0 \tag{7}$$

The association between variable $\mathbf{p}_i$ and variable $\mathbf{c}_j$ can be characterized as:

$$\mathbf{p}_j^T\mathbf{c}_j = 1, i = j$$
$$\mathbf{p}_j^T\mathbf{c}_j = 0, i \neq j \tag{8}$$

Therefore, matrices $\mathbf{X}$ and $\mathbf{Y}$ can be expressed as follows:

$$\mathbf{X} = \mathbf{U}\mathbf{C}^T + \mathbf{X}_0$$
$$\mathbf{Y} = \mathbf{U}\mathbf{E}^T + \mathbf{Y}_0 = \mathbf{X}\mathbf{P}\mathbf{E}^T + \mathbf{Y}_0 \tag{9}$$

Simultaneously, the aforementioned equation can demonstrate the correlation between $\mathbf{X}$ and $\mathbf{Y}$.

(6) Through additional calculations, it is possible to derive the reductive regression equation from $\mathbf{X}$ to $\mathbf{Y}$.

$$\hat{y} = a_1 x_1 + a_2 x_2 + \cdots + a_k x_k + \mathbf{Y_0} \tag{10}$$

The regression coefficient, a_i (where $i = 1, 2, ..., k$), represents the coefficient associated with each variable in the final reduction equation.

2.3.2. GA-SVR Model

SVR is an algorithm that has been developed based on the principles of support vector machine (SVM). The SVR model aims to achieve an optimal model by simultaneously minimizing the total fitting loss and maximizing the distance between the sample points and the hyperplane function. The SVR model is well-suited for analyzing small data sets due to its robust generalization capabilities, stability, and high accuracy in predicting outcomes. Therefore, it is deemed appropriate to employ the SVR model in this study to characterize the volatility of air passenger flows. The SVR model, which is fitted by a nonlinear function $f(x_i) = \omega^T x + b$, can be mathematically represented as follows:

$$\min \frac{1}{2}\|\omega\|^2 + C\sum_{r=1}^{m}(\xi_i + \xi_i^*)$$
$$s.t.\begin{cases} f(x_i) - y_i \leq \varepsilon + \xi_i \\ y_i - f(x_i) \leq \varepsilon + \xi_i^* \\ \xi_i > 0, \xi_i^* > 0, i = 1, 2, \cdots, m \end{cases} \tag{11}$$

Within this set, the penalty parameter is denoted as C, the tolerance parameter is represented by ε, and the relaxation variables are denoted as ξ_i and ξ_i^*. The utilization of the Lagrange multiplier technique is employed for solving the problem at hand. To accomplish this, non-negative Lagrange multipliers, namely α_i, α_i^*, β_i, and β_i^*, are introduced. Subsequently, the Lagrange function is formulated, and its partial derivatives with respect to ω, b, ξ_i, and ξ_i^* are equated to zero. By performing the aforementioned calculation, we have successfully derived the dual problem and the ultimate decision function for SVR.

$$\max \sum_{i=1}^{m} y_i(\beta_i^* - \beta_i) - \varepsilon(\beta_i^* + \beta_i) - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}(\beta_i^* - \beta_i)(\beta_j^* - \beta_j)x_i^T x_j$$
$$s.t.\begin{cases} \sum_{i=1}^{m}(\beta_i^* - \beta_i) = 0 \\ 0 \leq \beta_i^*, \beta_i \leq C \end{cases} \tag{12}$$

$$f(x_i) = \sum_{i=1}^{n}(\beta_i^* - \beta_i)x_i^T x_j + y_i + \varepsilon - \sum_{j=1}^{n}(\beta_j^* - \beta_j)x_j^T x_i \tag{13}$$

In equation (13), the variable n represents the quantity of support vectors. Given the solution to equation (13), it is observed that only the coefficient $(\beta_i^* - \beta_i)$ of the samples falling outside the interval band is non-zero. Consequently, these particular samples can be identified as support vectors, with their total count denoted as n.

The construction of an SVR model often encounters a challenge known as the linear inseparability problem. This means that a data set cannot be classified by a linear classifier (straight line or plane), which is quite common in practical applications. To address this issue, it becomes imperative to employ a technique that involves mapping the two-dimensional samples to a higher-dimensional space. By doing so, it becomes feasible to utilize a hyperplane for the purpose of partitioning. To streamline the inner product operation within the mapping space, it is widespread practice to incorporate the kernel function into the model. This study introduces the utilization of the radial basis kernel function for constructing the model, resulting in the final decision function being expressed in the form of the radial basis kernel function.

$$f(x_i) = \sum_{i=1}^{n}(\beta_i^* - \beta_i)k(x_i, x_j) + y_i + \varepsilon - \sum_{j=1}^{n}(\beta_j^* - \beta_j)k(x_i, x_j)$$

$$k(x_i, x_j) = \exp(-\frac{\left\| x_i - x_j \right\|^2}{2\sigma^2})$$

(14)

The σ represents the kernel parameter in the radial basis kernel function.

The performance of the SVR model can be enhanced by optimizing its parameters to achieve the most accurate fitting of the given data. Based on the aforementioned information, the model's adjustable parameters include the penalty parameter C, the tolerance parameter ε, and the kernel parameter σ. The utilization of GA is employed in this study for the purpose of parameter optimization. GA is capable of identifying the most optimal solution to a given problem through the simulation of the natural process of evolution, which involves selection, crossover, and mutation. The GA exhibits notable attributes such as rapid computational efficiency, exceptional precision, and seamless integration capabilities with other algorithms. Furthermore, given the limited sample size, the present study employs the cross-validation technique to enhance the robustness of the optimization procedure, thereby mitigating the risks of overfitting and underfitting. The parameters of the GA are configured according to the specifications presented in Table 1.

Table 1. Parameter setting of GA

Parameter	Value
Evaluation index	R^2
Population	20
Generations	50
Crossing rate	0.9
Mutation rate	0.1
Cross-validation fold	10

3. EMPIRICAL ANALYSIS

3.1. Influential Factor Selection and Variable Framework Development

The present analysis synthesizes the findings of Daldoul et al. (2016), Gundelfinger-Casar and Coto-Millán (2017), Yan and Chai (2017), Huang et al. (2013), Hanson et al. (2022), Tirtha et al. (2023). Collectively, these studies identify a total of fourteen primary factors that exert influence on the air passenger transportation market. Table 2 displays the variables that are pertinent to this study.

3.2. Sample Selection and Data Processing

In 2019, the Beijing-Shanghai route exhibited the highest volume of air passenger traffic within China. This study examines the effects of the COVID-19 epidemic on the civil aviation industry, focusing on the period from 2019 to the present. To conduct empirical testing and comparison, the study specifically selects the air passenger flow data between Beijing and Shanghai from 1999 to 2019. The dependent variable data used in this study are sourced from the China Civil Aviation Statistical Yearbook spanning the years 2000 to 2020. We use this dataset because this long series provides a sufficiently large set of actual operational data covering all the necessary routes to estimate passenger flows. At the same time, the dataset is compiled by the Development Planning Department of the China Civil Aviation Administration, which provides a certain guarantee of data quality. The original calculated data pertaining to economic and demographic characteristics are obtained from the Beijing Statistical Yearbook and Shanghai Statistical Yearbook, also covering the period from 2000 to 2020. The statistical yearbooks of all cities are compiled by the local government statistics office and their data quality is dependable.

In the present study, the initial dataset undergoes pre-processing, wherein a set of rules are employed.

(1) The random missing data is imputed through the use of interpolation techniques.
(2) The utilization of the random sample partition function is employed for the purpose of dividing the training set and the test set. The test set comprises 20% of the total dataset, and a fixed random seed of fifteen is used.
(3) The data is standardized using the Z-score standardization method. Z-score standardization is a common method of data processing that involves dividing the difference between the measured value and the mean by the standard deviation. Through this process, data of different magnitudes can be transformed into unified Z-scores for comparison.

3.3. Model Prediction Results

3.3.1. Prediction Results of PLS Model

Based on the modeling procedure of the PLS model, the initial step involves utilizing the training set to extract the principal components of the PLS model. This process results in the acquisition of three principal components. Subsequently, the dependent variables are employed for regression fitting and reduction calculation. Table 3 displays the coefficients and constants of the restored PLS regression equation. Table 4 displays the fitting and test results, along with their respective weights.

3.3.2. Prediction Results of GA-SVR Model

The initial step in the modeling and optimization of the GA-SVR model involves importing the training data and subsequently utilizing the GA to optimize the parameters. The modeling parameters that exhibit the highest fitness under cross-validation are chosen after undergoing numerous iterations. Through a series of extensive experiments, it has been determined that the penalty parameter C yields optimal fitness at a value of 5.61952194. Similarly, the tolerance parameter ε demonstrates optimal fitness at a value of 0.23896516, while the kernel parameter σ exhibits optimal fitness at a value of

Table 2. Variable names and calculations

Category	Variable			Unit	Label	Calculation method	Reference
Dependent variable	Market demand for air passenger transport			Person-times	PT	Total passenger traffic on routes in current year	None
Independent variable	Economic characteristics	City pairs' GDP		One hundred million yuan	GDP	$GDP^1 + GDP^2$	Yan and Chai (2017) Hanson et al. (2022)
		City pairs' added value of secondary industry		One hundred million yuan	$GDP2$	$GDP2^1 + GDP2^2$	
		City pairs' added value of tertiary industry		One hundred million yuan	$GDP3$	$GDP3^1 + GDP3^2$	
		City pairs' average growth rate of GDP		%	$GDPth$	$\frac{1}{2}\left(GDPth^1 + GDPth^2\right)$	
		City pairs' average value-added index of secondary industry		%	$GDP2th$	$\frac{1}{2}\left(GDP2th^1 + GDP2th^2\right)$	
		City pairs' average value-added index of tertiary industry		%	$GDP3th$	$\frac{1}{2}\left(GDP3th^1 + GDP3th^2\right)$	
		City pairs' proportion of added value of secondary industry		%	$GDP2/ GDP$	$\dfrac{GDP2^1 + GDP2^2}{GDP^1 + GDP^2}$	Huang et al. (2013)
		City pairs' proportion of added value of tertiary industry		%	$GDP3/ GDP$	$\dfrac{GDP3^1 + GDP3^2}{GDP^1 + GDP^2}$	
	Population and social characteristics	City pairs' resident population		Ten thousand people	Pop	$Pop^1 + Pop^2$	Tirtha et al. (2023)
		City pairs' average per capita GDP		yuan	$GDPper$	$\frac{1}{2}\left(\dfrac{GDP^1}{Pop^1} + \dfrac{GDP^2}{Pop^2}\right)$	Gundelfinger-Casar and Coto-Millán (2017)
		City pairs' retail sales of consumer goods		One hundred million yuan	$Sale$	$Sale^1 + Sale^2$	Huang et al. (2013)
		City pairs' per capita disposable income		yuan	$Wage$	$Wage^1 + Wage^2$	Hanson et al. (2022) Tirtha et al. (2023)
		City pairs' employees at the end of the year		Ten thousand people	Emp	$Emp^1 + Emp^2$	Daldoul et al. (2016)
		City pairs' average resident population density		Person/km²	$Density$	$\frac{1}{2}\left(Density^1 + Density^2\right)$	

Note: The variable superscripts 1 and 2 are used to denote two distinct cities within a given city pair.

Table 3. Results of PLS model

Variables	Value
GDP	0.0013
GDP2	0.0082
GDP3	0.0010
GDPth	103 820.7660
GDP2th	24 655.6813
GDP3th	30 515.0881
GDP2/GDP	-43 491.4762
GDP3/GDP	-46 467.9169
Pop	554.3408
GDPper	7.3564
Sale	0.0027
Wage	2.1409
Emp	-430.0975
Density	1 013.4680
Cons	2 256 396.5625

Table 4. Single model results, weights, and variable weight combined model results

Model	PLS		GA-SVR		Variable weight combined model
	PT	Weight	PT	Weight	PT
Fitting value	5 125 826	0.89484	5 387 501	0.10516	5 153 343
	2 725 188	0.21365	3 170 962	0.78635	3 075 722
	6 635 386	0.58127	6 699 746	0.41873	6 662 336
	6 780 267	0.45827	6 771 230	0.54173	6 775 372
	3 150 692	0.69123	3 302 510	0.30877	3 197 569
	7 366 541	0.89116	7 089 891	0.10884	7 336 430
	6 775 343	0.54887	6 735 605	0.45113	6 757 416
	2 671 950	0.99359	2 845 498	0.00641	2 673 062
	4 340 908	0.99927	4 088 119	0.00073	4 340 723
	7 738 793	0.56925	7 234 201	0.43075	7 521 440
	6 458 389	0.71189	6 262 785	0.28811	6 402 034
	5 044 701	0.85588	4 753 396	0.14412	5 002 718
	2 266 327	0.96536	2 611 821	0.03464	2 278 295
	3 698 585	0.72978	3 589 258	0.27022	3 669 042
	6 435 067	0.97923	6 461 195	0.02077	6 435 610
	5 807 115	0.99820	5 362 538	0.00180	5 806 314
Test value	5 398 102	0.07955	5 664 360	0.92045	5 643 179
	6 324 054	0.94545	6 484 784	0.05455	6 332 822
	2 652 412	0.88711	3 342 230	0.11289	2 730 286
	8 226 784	0.07441	7 816 708	0.92559	7 847 222
	8 649 517	0.03988	7 799 476	0.96012	7 833 375

0.00817128. The graph depicting the fitness curve is presented in Fig. 2. The SVR model was constructed using the aforementioned parameters for the purposes of fitting and testing. The outcomes and corresponding weight values are presented in Table 4.

3.3.3. Prediction Results of Variable Weight Combined Model

Based on the procedural guidelines of the variable weight combined model, the initial step involves determining the weights of each individual model at each given instance. This is accomplished through the utilization of the weight calculation formula in conjunction with the outcomes of each respective single model. Subsequently, the variable weight combined model is evaluated by computing the fitting and test results based on the assigned weights. These results are presented in Table 4. Fig. 3 shows the comparison between the predicted value and the actual value of the PLS, GA-SVR, and variable weight combined model, where the blue line indicates that the horizontal axis is equal to the vertical axis, and the five-pointed star indicates the relative relationship between the predicted value and the actual value. When the pentagram is closer to the blue line, the gap between the predicted value and the actual value is smaller. From the visualization results, it can be seen that the prediction effect of the variable weight combination prediction model is better.

3.3.4. Model Comparison

This study compares the coefficient of determination (referred to as R^2) and the mean absolute percentage error (referred to as MAPE) of each model to assess the level of agreement between the predicted and actual values. A higher R^2 indicates a greater level of concordance between the predicted value and the observed value. A decrease in the MAPE corresponds to a decrease in the level of error. Table 5 displays the R^2 and MAPE values for each model during both the training and testing stages. The variable weight combined model exhibits a higher R^2 value compared to alternative models, indicating a stronger correlation between the predicted and actual values. Additionally,
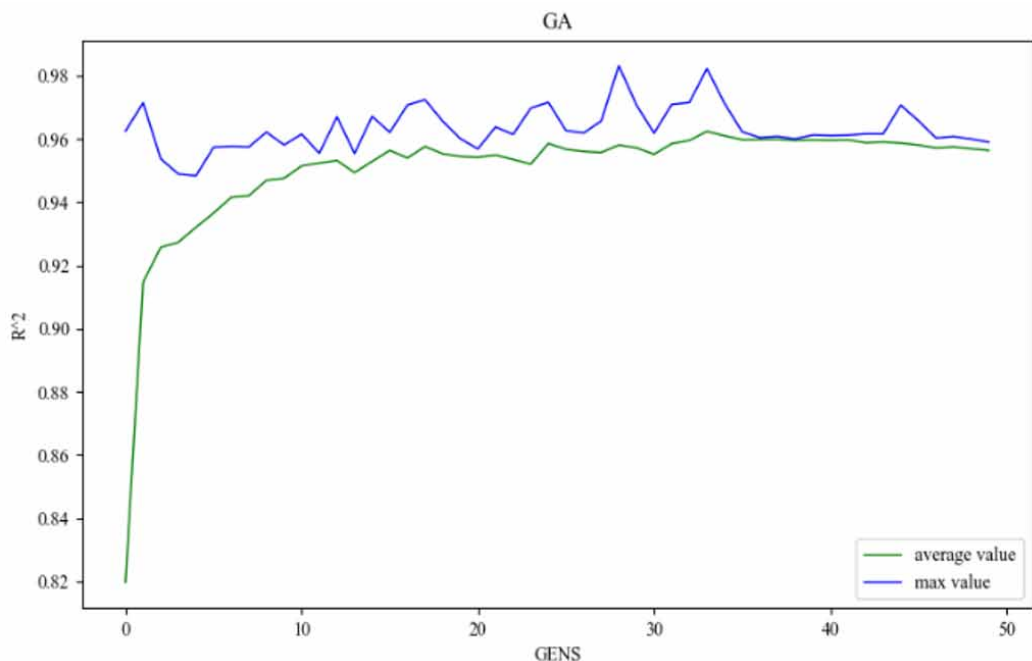
Figure 2. Fitness curve

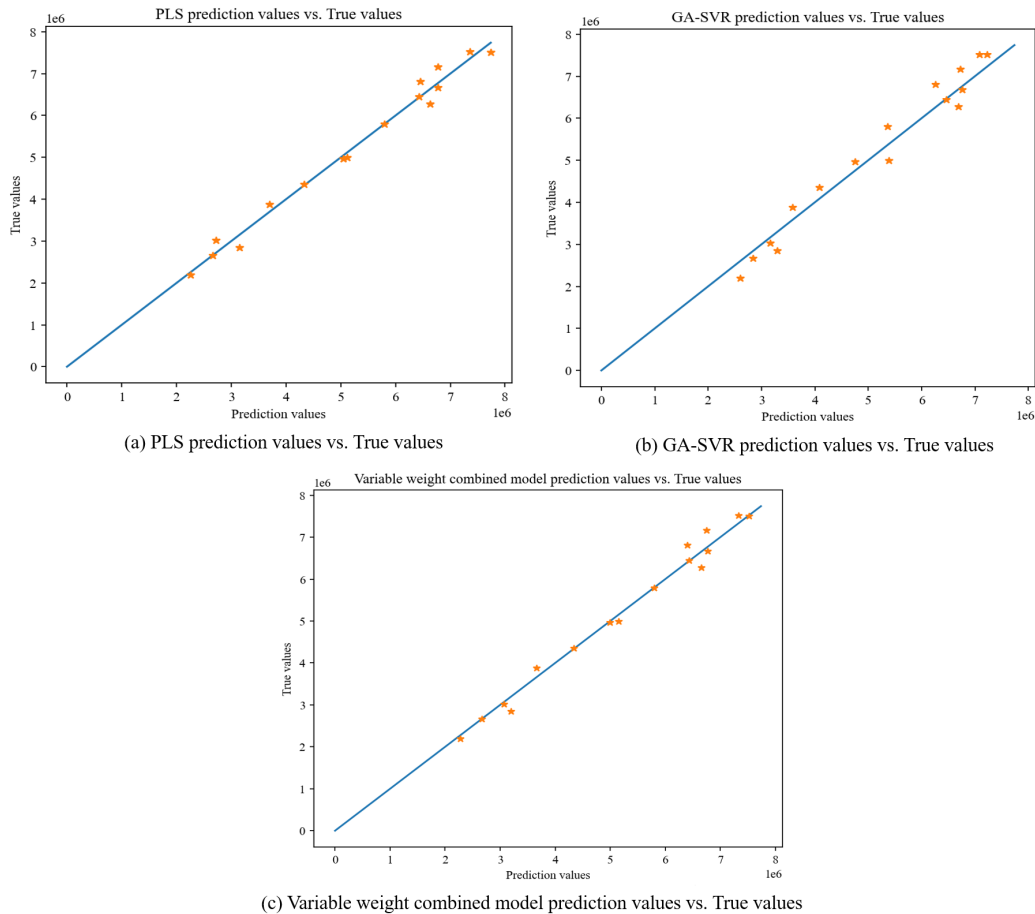**Figure 3. Comparison between the prediction effect of the model and the real value**

(a) PLS prediction values vs. True values

(b) GA-SVR prediction values vs. True values

(c) Variable weight combined model prediction values vs. True values

Table 5. Evaluation index results of models

Models	Index	Evaluation of fitting results	Evaluation of test results
PLS	MAPE	0.03575	0.05273
	R^2	0.98566	0.96672
GA-SVR	MAPE	0.07016	0.04972
	R^2	0.96265	0.98190
Fixed weight combined model	MAPE	0.03582	0.05098
	R^2	0.98566	0.95754
Variable weight combined model	MAPE	0.03180	0.01687
	R^2	0.98577	0.99746

the variable weight combined model demonstrates a lower MAPE in comparison to other models, suggesting a higher level of accuracy in predicting the target variable. Simultaneously, the variable weight combined model exhibits superior generalization capabilities in comparison to the fixed weight combined model. Hence, this study has effectively achieved model optimization and developed an enhanced mathematical model for forecasting air passenger flows. In addition, the results also show that the information extraction of the comprehensive development of social economy from the two aspects of linear development trend and nonlinear fluctuation law can effectively improve the fitting effect of the final model and achieve more accurate prediction.

3.4. Predicting Future Trends

3.4.1. Predicting Future Values of Influential Factors

To accurately predict the future value, it is imperative to effectively forecast the factors that exert influence. The ARIMA model, which stands for autoregressive integrated moving average, is employed to identify the hyperparametric model that minimizes the information criterion. The prediction model is depicted in Table 6.

3.4.2. Predicting the Air Passenger Flows

The influencing factors' predicted values are inputted into two separate models for computation, resulting in the predicted values of each individual model. The weight is determined using formula (2), and the future value projected by the composite model incorporating the predicted values from each individual model can be obtained by combining them. Considering the significant ramifications of the COVID-19 pandemic on the growth and functioning of the civil aviation sector between 2020 and 2022, there was a substantial decline in business data. However, with the introduction of the "New Ten Principles" towards the end of 2022 and the complete resumption of China's international passenger flights on 8 January 2023, the civil aviation industry is expected to gradually recover and regain its pre-pandemic (2019) level of development starting from 2023. Hence, this research employs the year 2023 as the subsequent year following 2019 to acquire the projected value. The outcomes are presented in Table 7.

Table 6. Models used to predict future values of factors

Variables	Models
GDP	ARIMA (0,2,0)
GDP2	ARIMA (3,2,2)
GDP3	ARIMA (0,2,0)
GDPth	ARIMA (2,2,0)
GDP2th	ARIMA (3,2,0)
GDP3th	ARIMA (3,2,0)
GDP2/GDP	ARIMA (2,2,2)
GDP3/GDP	ARIMA (0,2,3)
Pop	ARIMA (0,2,0)
GDPper	ARIMA (0,2,0)
Sale	ARIMA (0,2,0)
Wage	ARIMA (0,2,0)
Emp	ARIMA (3,2,0)
Density	ARIMA (0,2,1)

Table 7. Results of variable weight combined model

Year	Beijing-Shanghai *PT*
2023	7 788 294
2024	8 410 047
2025	8 998 242
2026	9 523 074
2027	10 199 264

4. ANALYSIS OF PREDICTING RESULTS

4.1. Predicting Results of Air Passenger Flows Among Major Hub Cities in China

This study aims to enhance the utilization of the model in predicting and comparing air passenger flows among major hub cities in China. Based on the data sourced from the China Civil Aviation Statistical Yearbook for the year 2020, an analysis was conducted on the passenger flows within the major multi-airport systems of mainland China, specifically focusing on the six hub cities. China represents a considerable proportion, specifically 10%, of the overall air passenger transport sector. Consequently, the market value associated with these hub cities is notably elevated. This study utilizes a predictive model to estimate the air passenger flows among the designated hub cities. The primary source of data utilized in this study is derived from the China Civil Aviation Statistical Yearbook, as well as the statistical yearbooks of various cities, spanning the period from 2000 to 2020. The models have successfully fit the data from these thirteen routes, and the corresponding predicted outcomes are presented in Table 8. Overall, China's inter-city hub market will exceed three hundred million passengers in the next five years, with an average annual growth rate of 6.51%.

Table 8. Prediction results among hub cities

City pairs	2023~2027		2027		
	PT	Average growth rate	*PT*	Increase compared to 2019	Growth rate compared to 2019
Beijing-Shanghai	44 918 921	5.19%	10 199 264	2 255 879	28.40%
Shanghai-Shenzhen	39 467 594	8.03%	9 040 683	2 887 156	46.92%
Beijing-Chengdu	35 441 977	10.28%	7 981 505	3 033 686	61.31%
Shanghai-Guangzhou	35 028 854	5.19%	7 900 572	1 758 240	28.62%
Beijing-Guangzhou	31 170 293	4.71%	6 879 694	1 411 197	25.81%
Beijing-Shenzhen	28 853 639	5.27%	6 356 499	1 438 718	29.26%
Shanghai-Chengdu	24 484 939	6.44%	5 520 367	1 479 880	36.63%
Shanghai-Chongqing	20 843 788	8.55%	4 924 631	1 654 852	50.61%
Chengdu-Guangzhou	18 301 594	5.05%	3 925 851	851 533	27.70%
Chengdu-Shenzhen	17 180 089	6.33%	3 901 854	1 024 948	35.63%
Chongqing-Guangzhou	16 500 232	9.15%	3 722 805	1 308 136	54.17%
Chongqing-Shenzhen	15 342 351	3.84%	3 141 508	477 256	17.91%
Beijing-Chongqing	15 220 483	6.61%	3 323 044	902 842	37.30%
Total	342 754 754	6.51%	76 818 277	20 484 323	36.94%

4.2. Management Decision Analysis and Debate

To optimize the efficiency of construction and operation, as well as minimize resource mismatch and waste, it is imperative to align the diverse market demands with the varying capacity inputs of aviation departments and government planning. Among the predicted outcomes, the air passenger volume between Shanghai and Beijing is the highest, surpassing forty million individuals. The projected air passenger volumes for the Shanghai-Shenzhen, Beijing-Chengdu, and Shanghai-Guangzhou routes all exceed thirty-five million individuals. Additionally, the projected air passenger volumes for the Beijing-Guangzhou, Beijing-Shenzhen, Shanghai-Chengdu, and Shanghai-Chongqing routes all surpass twenty million individuals. Lastly, the projected air passenger volumes for the remaining five city pairs all exceed ten million individuals. In relation to the management of airspace resources and shipping capacity, it may be advisable for the government to contemplate elevating the anticipated air passenger flow between the specified city pairs that exhibit substantial predictability. This strategic approach aims to maximize the utilization of demand, establish routes of superior value and quality, and enhance the overall standard of the national route network.

By conducting comparative analysis of projected air passenger flows and the average annual growth rate for the years 2019 and 2027, it is possible to ascertain the strength of the growth momentum. Based on the product life cycle theory, the presence of robust growth potential indicates that the trajectory is currently situated within the growth phase of market expansion. Consequently, it is anticipated that the volume of air passenger traffic will sustain continuous growth over an extended period. Diverse growth expectations necessitate varying allocations of long-term investment between the government and enterprises. These investments may include government development planning, infrastructure construction, airline brand development, and the establishment of airline bases, among others. According to the projected outcomes, the air passenger traffic between Beijing and Chengdu exhibits the most substantial annual growth rate (10.28%) and growth volume (3,033,687). The mean yearly growth rate of the Shanghai-Chongqing and Chongqing-Guangzhou routes exceeds 8.5%. Therefore, the routes exhibit greater growth prospects, and their long-term investment and development yield more advantageous outcomes in enhancing the sustainable development capacity of the national route network quality.

5. CONCLUSION

This study utilizes the principle of air passenger flow development between cities to establish a variable weight combined model, which integrates the PLS model and the GA-SVR model. The model comprehensively describes the linear development trend and fluctuation law of economy and society to predict the future air passenger traffic volume. The results show that the model constructed in this study is better than the weighted combination forecasting model and other traditional single models.

The empirical testing of the variable weight combined model is conducted using the Beijing-Shanghai air passenger flow data. The results indicate that both the R^2 and MAPE metrics are significantly improved compared to the individual models and fixed weight combined model. Hence, the optimization of the model has been effectively achieved, resulting in the development of an improved mathematical model for forecasting air passenger flows. This accomplishment offers a novel approach for predicting market demand in the air passenger industry. This shows that the information extraction of the comprehensive development of social economy from the two aspects of linear development trend and nonlinear fluctuation law can effectively improve the fitting effect of the final model and achieve more accurate prediction. This result provides a new way of thinking for the demand forecast of airline passenger transport market.

Secondly, this study applies the variable weight combination forecasting model to thirteen routes among the major hub cities in China and obtains the market demand forecast and expected growth

rate of each route. And finally, the overall expected development of the air transportation market among the hub cities in mainland China is obtained. The results show that the passenger flow in the air transportation market between the hub cities in mainland China has an average annual growth rate of 6.51% from 2023 to 2027, with a total passenger flow of more than three hundred million. According to the analysis of this study, this is due to the fact that during the novel coronavirus epidemic, mainland China's economic growth trend remained positive and did not show negative growth, although its growth rate decreased and its aviation market continued to recover, so the recovery of Chinese passenger volume is relatively not surprising.

Then, this study analyzes the route-based market segmentation, which can help to reveal more information about the internal structure of the air passenger transport market among China's hub cities. The study employs a variable weight combined model to forecast the air passenger flows among six major hub cities in mainland China. The predicted value and anticipated growth rate are acquired. Based on projected air passenger flows, it is anticipated that the routes connecting Shanghai-Beijing, Shanghai-Shenzhen, Beijing-Chengdu, and Shanghai-Guangzhou will experience higher volumes of air passenger traffic from 2023 to 2027. Subsequently, the routes of Beijing-Guangzhou, Beijing-Shenzhen, Shanghai-Chengdu, and Shanghai-Chongqing are expected to exhibit moderate air passenger flow volumes. In terms of average annual growth rate, the Beijing-Chengdu route is projected to have the highest rate of growth in air passenger flow, followed by Shanghai-Chongqing and Chongqing-Guangzhou. Hence, considering the routes exhibiting robust market demand and growth potential, it is advisable for the government, airline companies, and airports to allocate additional resources towards thoroughly investigating the market demand within these cities. This strategic approach will facilitate the advancement of a high-caliber and sustainable route network.

Finally, it would be useful to explore whether and to what extent our findings are applicable to other countries. For example, the American aviation market, which is also affected by novel coronaviruses, is also influenced by the overall economic and social environment of the United States, and the forecast of passenger flows on American routes is also applicable to the model built in this study.

6. LIMITATIONS AND FUTURE RESEARCH

This study has several limitations that provide opportunities for further research and improvement. First, further research and exploration of passenger flow forecasts for air transportation markets in other countries may confirm and extend our results. Second, there may be differences in air passenger flow affected by the linear development trend and fluctuation law of economic society. Future research can explore the different degrees of impact of linear development trends and fluctuation law on air transport passenger flow in different situations. Third, although our samples come from several routes, they are only located in one country. Therefore, our samples may have limitations and need to be studied on a larger scale, including cross-regional, market size and hub city level routes, which can further deepen and verify our findings.

Several issues related to our research have not yet been explored. First, the study can further discuss the demand for diverse types of aircraft in the future aviation market based on the forecast of air traffic passenger flows. The experiences and lessons learned from these studies can help airlines and stakeholders to better understand how passenger flows lead to investment and financing capacity and operating costs of companies. Second, it is useful to conduct some more detailed studies, such as discussing and measuring the importance of various economic and social factors in the forecast, which may help airlines, airports, and stakeholders to judge the development trend of air passenger flow quickly and accurately by observing the changes

of decisive factors in the business environment when economic fluctuations occur in the future. Thus, it is especially important to make more flexible strategic decisions and value assessment in the post-epidemic era with high uncertainty.

ACKNOWLEDGMENT

REFERENCES

Chen, C., & Li, C. (2022). Passenger Traffic Forecast of Civil Aviation Based on PSO-ELM Algorithm. *Kongzhi Gongcheng*, *29*(11), 2042–2047. doi:10.14107/j.cnki.kzgc.20200682

Daldoul, M., Jarboui, S., & Dakhlaoui, A. (2016). Public transport demand: Dynamic panel model analysis. *Transportation*, *43*(3), 491–505. doi:10.1007/s11116-015-9586-1

Deng, S., Zhou, L., Wang, F., & Liu, Z. (2022). XGBoost-LSTM Combination Model with variable weight for Dam Deformation Prediction and Its Application. *Changjiang Kexueyuan Yuanbao*, *39*(10), 72–79. doi:10.11988/ckyyb.20210641

Gundelfinger-Casar, J., & Coto-Millán, P. (2017). Intermodal competition between high-speed rail and air transport in Spain. *Utilities Policy*, *47*, 12–17. doi:10.1016/j.jup.2017.06.001

Hanson, D., Toru Delibasi, T., Gatti, M., & Cohen, S. (2022). 2022/01/01/). How do changes in economic activity affect air passenger traffic? The use of state-dependent income elasticities to improve aviation forecasts. *Journal of Air Transport Management*, *98*, 102147. doi:10.1016/j.jairtraman.2021.102147

Huang, B., Lin, J., Zheng, X., & Fang, X. (2013). The Prediction for Civil Airport Passenger Throughput of Based on Multiple Linear Regression Analysis. *Practice and Understanding of Mathematics*, *43*(04), 172–178.

Kumbure, M. M., Lohrmann, C., Luukka, P., & Porras, J. (2022). Machine learning techniques and data for stock market forecasting: A literature review. *Expert Systems with Applications*, *116659*, 116659. doi:10.1016/j.eswa.2022.116659

Li, L., Fu, X., Lei, Z., Wang, K., & Seno-Alday, S. (2020, March). Quantifying market potential for sustainable air connectivity improvement - A counterfactual analysis based on augmented gravity model estimates [Article]. *Research in Transportation Business & Management*, *34*(13), 100425. doi:10.1016/j.rtbm.2019.100425

Liang, X., Qiao, H., Wang, S., & Zhang, X. (2017). Integrated Forecasting Model for Air Passenger Traffic in China Based on Singular Spectrum Analysis. *System Engineering—Theory & Practice, 37*(06), 1479-1488. 10.16381/j.cnki.issn1003-207x.2019.1100

Pereira, J., Braga, V., Correia, A., & Salamzadeh, A. (2021). Unboxing organisational complexity: How does it affect business performance during the COVID-19 pandemic? *Journal of Entrepreneurship and Public Policy*, *10*(3), 424–444. doi:10.1108/JEPP-06-2021-0070

Shahul Hameed, N. S., Salamzadeh, Y., Abdul Rahim, N. F., & Salamzadeh, A. (2022). The impact of business process reengineering on organizational performance during the coronavirus pandemic: moderating role of strategic thinking. *foresight, 24*(5), 637-655. 10.1108/FS-02-2021-0036

Shen, J., Zeng, X., & Wu, G. (2019). Research on Dynamics Model of China's Civil Aviation Passenger Transport Market Demand Forecast. *Journal of East China Jiaotong University*, *36*(04), 57–66. doi:10.16749/j.cnki.jecjtu.2019.04.008

Tirtha, S. D., Bhowmik, T., & Eluru, N. (2023). 2023/01/01/). Understanding the factors affecting airport level demand (arrivals and departures) using a novel modeling approach. *Journal of Air Transport Management*, *106*, 102320. doi:10.1016/j.jairtraman.2022.102320

Truong, D. (2021). 2021/09/01/). Estimating the impact of COVID-19 on air travel in the medium and long term using neural network and Monte Carlo simulation. *Journal of Air Transport Management*, *96*, 102126. doi:10.1016/j.jairtraman.2021.102126 PMID:36569042

Wierzbinski, B., Zaborek, P., Wosiek, M., & Surmacz, T. (2023). 04/20). Knowledge management orientation as a driver of competitive performance: Evidence from Polish SMEs in the aviation industry. *Entrepreneurial Business and Economics Review*, *11*(1), 93–112. doi:10.15678/EBER.2023.110105

Wu, Z., & Xiong, J. (2021). 04/01). A Novel Task-Scheduling Algorithm of Cloud Computing Based on Particle Swarm Optimization. *International Journal of Gaming and Computer-Mediated Simulations*, *13*(2), 1–15. doi:10.4018/IJGCMS.2021040101

Yan, H., & Chai, H. (2017). Evaluation of the influencing factors on general aviation tourism industry of xi'an based on AHP and fuzzy comprehensive evaluation method. *2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*. IEEE.

Zhang, L., Yang, H., Wang, K., Bian, L., & Zhang, X. (2021). 2021/08/01/). The impact of COVID-19 on airline passenger travel behavior: An exploratory analysis on the Chinese aviation market. *Journal of Air Transport Management*, *95*, 102084. doi:10.1016/j.jairtraman.2021.102084 PMID:36568571

Zhang, P. (2018). Comparison of application of variable weight and fixed weight in combination forecasting. *Statistics & Decisions*, *34*(17), 80–82. doi:10.13546/j.cnki.tjyjc.2018.17.019

Zhang, Y., Cao, K., & Dong, W. (2021). Research on improvement and optimisation of modelling method of China's civil aircraft market demand forecast model. *Aeronautical Journal*, *125*(1289), 1157–1191. doi:10.1017/aer.2021.16