ISSN 2616-5775 Vol. 9, Issue 4: 56-60, DOI: 10.25236/AJCIS.2026.090407

Integrated Prediction of Multidimensional Civil Aviation Operational Indicators: A Multi-Task Deep Neural Network Approach

Lu Qinghui[1,a,*], Liu Yi[1,b], Zhang Yun[1,c]

[1] *University of Science and Technology Liaoning, Anshan, China*
[a] *1229108840@qq.com*, [b] *3061477562@qq.com*, [c] *2684807100@qq.com*
[*] *Corresponding author*

Abstract: *Aiming at the coupling of the three key indicators—flight delay, revenue, and carbon emissions in civil aviation, as well as the problems that traditional single-task prediction fails to mine internal correlations and lacks sufficient accuracy, this paper proposes an integrated prediction method based on multi-task deep neural networks. Based on the PIA 2026 aviation dataset, we complete data preprocessing and feature engineering, then construct and compare the random forest single-task model and the multi-task deep neural network model. The results show that the multi-task model can effectively capture the correlations among indicators, achieve higher accuracy in delay prediction, perform synchronous prediction of the three indicators, and possess stronger generalization ability. It can provide decision support for airline operation and low-carbon development, and improve the theoretical system of intelligent prediction in civil aviation. This study provides a feasible intelligent decision support scheme for the coordinated optimization of safety, benefit, and green development in civil aviation operations.*

Keywords: *Civil Aviation Operation Indicators; Integrated Prediction; Multi-Task Deep Neural Network; Flight Delay; Carbon Emission*

1. Introduction

In the digital transformation of civil aviation, flight delay, revenue, and carbon emissions constitute three mutually coupled core operational indicators that jointly affect airline efficiency, economic benefits, and low-carbon development[1]. Due to their inherent correlations and complex interactions, traditional single-task prediction models often fail to effectively capture the underlying relationships among these indicators, resulting in insufficient prediction accuracy and weak generalization performance[2]. To address these limitations, this paper proposes a multi-task deep learning–based prediction framework, constructs a single-task random forest model and a multi-task deep neural network (MTDNN) model, respectively, and conducts a comprehensive comparative analysis to verify the effectiveness of the proposed method.

2. Related Work

In recent years, flight delay prediction, revenue estimation, and carbon emission calculation have become research hotspots in intelligent civil aviation. Traditional methods mainly rely on statistical models, machine learning models such as random forest, support vector machine, and LightGBM to complete single-task prediction. Although these methods achieve certain effects in specific indicators, they ignore the coupling relationship among delay, revenue, and carbon emissions. With the development of deep learning, multi-task learning has been gradually applied to aviation operation prediction. By sharing underlying feature representation, multi-task models can mine potential correlations among multiple tasks and improve overall prediction performance and generalization ability. However, most existing studies focus on single-dimensional prediction such as delay or carbon emission, and few studies realize the integrated prediction of the three core indicators of civil aviation operation. Therefore, this paper constructs a multi-task deep neural network to realize synchronous prediction and correlation mining of flight delay, revenue, and carbon emissions.

ISSN 2616-5775 Vol. 9, Issue 4: 56-60, DOI: 10.25236/AJCIS.2026.090407

3. Datasets and Data Preprocessing

3.1. Data Exploration and Feature Analysis

The experimental dataset consists of 800 complete flight records and 22 original features covering major civil aviation operation scenarios. Data preprocessing inspection shows that no missing values, outliers, or abnormal entries exist in the dataset, demonstrating high data quality and reliability for model training and validation. Among all features, 6 categorical variables are designed in accordance with real-world civil aviation business rules and operational characteristics[3]. Statistical analysis reveals that international routes account for 88.25% of the total samples, and the distribution of aircraft types is relatively balanced, which helps improve the generalization ability of the prediction model[4].These features include flight basic information, aircraft parameters, route attributes, operational data, and market-related indicators, which cover the main influencing factors of delay, revenue, and carbon emissions.

3.1.1. Analysis of Flight Delay Features

As depicted in Figure 1, the distribution of flight delay duration presents a noticeable multi-peak pattern. The delay values are mainly concentrated within the interval of 0–250 minutes, with a prominent peak appearing at around 170 minutes. Further statistical analysis indicates that severe delay events account for 49.75% of the total samples, while moderate delays constitute 37.5%, reflecting a high proportion of non-negligible medium and long-duration delays in the dataset.

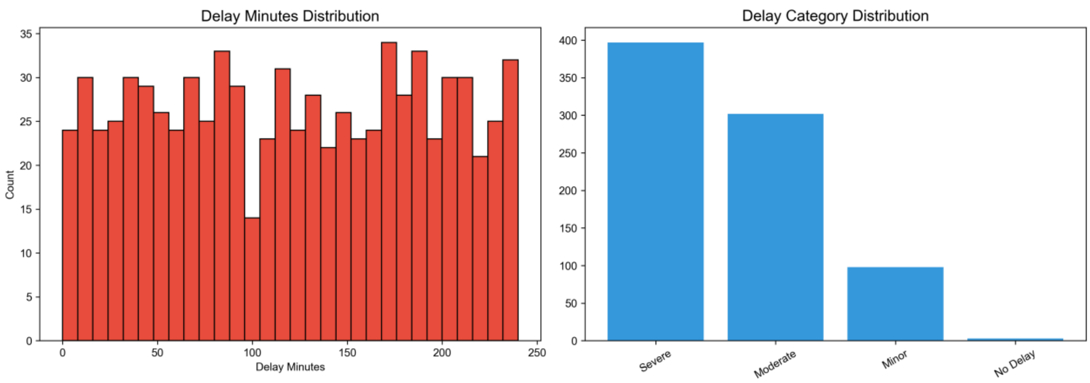

Figure 1: Distribution Characteristics of Flight Delays

3.1.2. Correlation Analysis between Aircraft Type and Carbon Emissions

As illustrated in Figure 2, noticeable differences exist in average carbon emissions across different aircraft types. Among them, the Boeing 737 registers the highest average carbon emission level, followed by the ATR 72 and Boeing 777. In contrast, the Airbus A320 exhibits the lowest average carbon emissions, indicating distinct emission characteristics associated with different aircraft models in civil aviation operations.

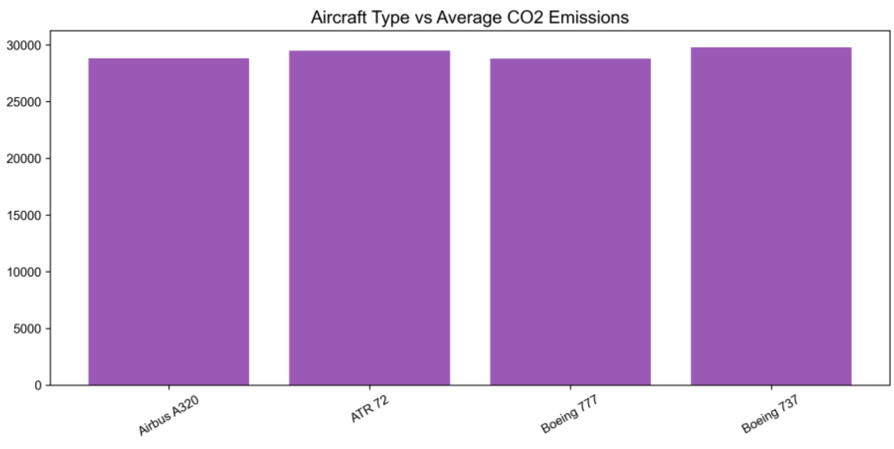

Figure 2: Correlation between Aircraft Type and Average Carbon Emissions

ISSN 2616-5775 Vol. 9, Issue 4: 56-60, DOI: 10.25236/AJCIS.2026.090407

3.1.3. Correlation Analysis between Load Factor and Revenue

As illustrated in Figure 3, a pronounced positive correlation can be observed between passenger load factor and route-level revenue. Specifically, intervals with higher load factor values correspond to a concentrated distribution of high revenue levels, which reflects that the occupancy rate of flights serves as a key driving factor for improving route revenue performance.[5]

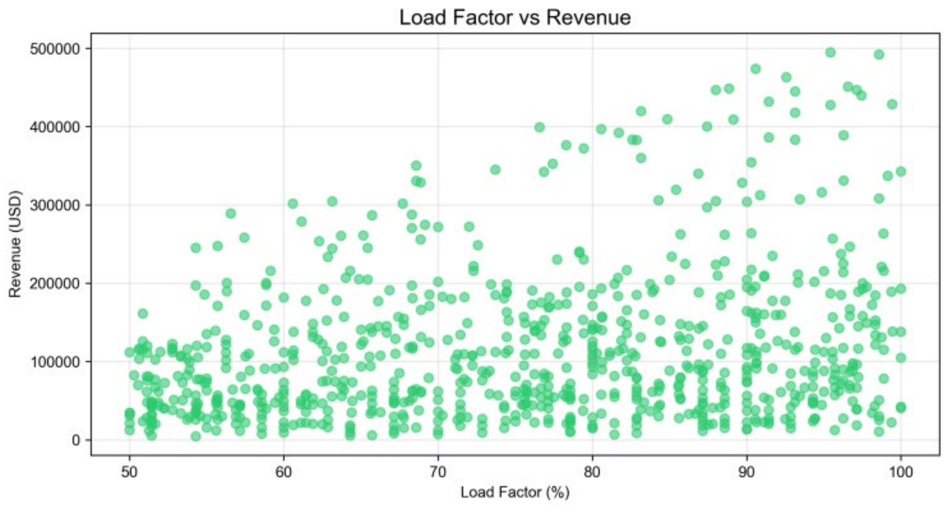

Figure 3: Scatter Plot of Load Factor vs. Route Revenue

3.2. Implementation of Data Preprocessing

Based on sufficient data exploration and statistical analysis, this study carries out systematic standardized preprocessing on the aviation dataset. Specifically, 10 representative numerical features are carefully selected as model inputs. To eliminate the adverse effects of inconsistent feature magnitudes and improve training stability, the Z-score standardization method is applied to normalize all numerical features. Meanwhile, all categorical features are properly integrated, encoded, and processed to form a unified feature structure for subsequent model construction and training[6].

4. Feature Engineering and Model Construction

4.1. Design and Implementation of Feature Engineering

To further enhance the expressive capacity of input features and enable the model to mine implicit patterns hidden in civil aviation operation data, this study additionally constructs three meaningful derived features based on domain knowledge and statistical characteristics. The three derived features include flight utilization ratio, seat efficiency, and route carbon emission intensity, which effectively enhance the representation of operational and environmental characteristics. Simultaneously, five categorical variables related to flight operations are converted into numerical vectors through one-hot encoding, so that discrete attribute information can be effectively recognized and utilized by deep learning models. Following a series of standardized procedures including feature screening, correlation analysis, and redundant information elimination, a total of 37 highly representative and informative core features are systematically integrated and determined as the final input feature set for formal model training and prediction.

4.2. Prediction Model Construction

Based on the PIA 2026 dataset, a single-task random forest model is constructed. Meanwhile, a multi-task deep neural network (MTDNN) is established, which takes 37-dimensional features as input, adopts three fully connected layers and three independent output heads, and uses the Adam optimizer and MSE joint loss function to achieve collaborative optimization[7].The MTDNN model takes the 37-dimensional fused features as input and passes through three fully connected hidden layers for feature abstraction. Each hidden layer adopts a linear transformation and activation function to enhance

ISSN 2616-5775 Vol. 9, Issue 4: 56-60, DOI: 10.25236/AJCIS.2026.090407

nonlinear expression ability. The model is designed with three independent output heads, which are respectively used for delay duration prediction, revenue regression, and carbon emission regression. The joint loss function is composed of the mean square error (MSE) of each task, so as to realize multi-task collaborative optimization. The Adam optimizer is used to accelerate convergence and improve training stability. For the single-task random forest model, the number of decision trees is set to 100, the maximum depth is limited to avoid overfitting, and other parameters adopt default optimized configurations to ensure fair comparison with the multi-task model.

4.3. Model Training Strategy

The dataset is partitioned into a training set consisting of 640 samples and a test set of 160 samples using a stratified random splitting strategy at an 8:2 ratio. During model training, the maximum number of epochs is set to 30 and the batch size to 8. In addition, an early stopping strategy is employed to prevent overfitting and improve training efficiency, while a fixed random seed is set to ensure experimental repeatability and result stability.

Figure 4 illustrates the training loss curve of the proposed MTDNN model. It can be observed that both the training loss and test loss gradually decrease and tend to stabilize after approximately 25 training epochs, with no obvious fluctuation or rebound phenomenon throughout the iteration process. This stable convergence trend fully verifies that the model has no overfitting and achieves reliable and robust fitting performance.

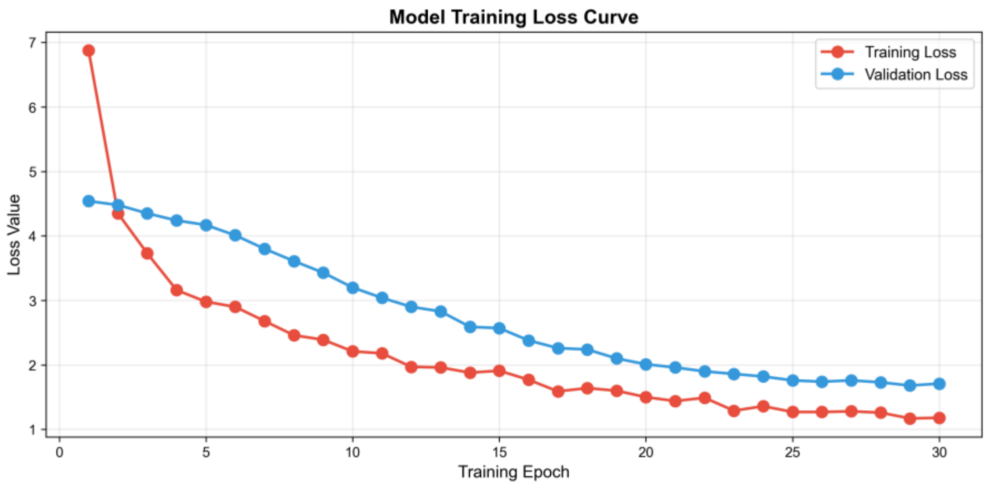

Figure 4: Training loss curve of the multi-task model

5. Experimental Results and Analysis

5.1. Evaluation Metrics

To quantitatively evaluate the prediction performance of the single-task and multi-task models, two widely used regression evaluation indicators are adopted in this paper: the coefficient of determination R^2 and the mean absolute error (MAE). R^2 reflects the degree of fitting between predicted values and real values; a value closer to 1 indicates better fitting effect. MAE represents the average absolute deviation between predicted and actual values, which directly reflects the prediction error level. The joint use of R^2 and MAE can comprehensively and objectively measure model accuracy, stability, and generalization ability.

5.2. Model Experimental Results

For the single-task random forest model, the training set achieved an R2 value of 0.68 and an MAE of 42.36 minutes, while the test set obtained an R2 of −0.12 and an MAE of 68.72 minutes; this substantial discrepancy between training and test performance reveals serious overfitting and poor generalization ability. In contrast, the multi-task deep neural network showed stable convergence in both training and testing losses with no evident overfitting, demonstrating significantly stronger generalization and more robust prediction performance[8].

ISSN 2616-5775 Vol. 9, Issue 4: 56-60, DOI: 10.25236/AJCIS.2026.090407

5.3. Model Comparative Analysis

The comparative experimental results demonstrate that the proposed multi-task MTDNN model achieves a test set R^2 value of 0.08 and an MAE of 59.45 minutes in flight delay duration prediction. In contrast to the single-task random forest model, which yields an R^2 of −0.12 and an MAE of 68.72 minutes on the same test set, the MTDNN model presents a substantial improvement in overall prediction accuracy and reliability. Furthermore, the multi-task learning framework supports the simultaneous prediction of flight delay, operational revenue, and carbon emissions, which effectively captures the inherent coupling relationships among these critical aviation indicators. The experimental results fully prove that the multi-task learning framework can share feature representation among related tasks, reduce prediction variance, and effectively improve generalization performance in civil aviation multi-index integrated prediction.

6. Conclusion

Based on the PIA 2026 aviation dataset, this paper establishes an integrated multi-task prediction framework using advanced deep learning techniques. The experimental results comprehensively demonstrate that the proposed multi-task model achieves higher prediction accuracy and better generalization performance than the traditional single-task random forest model[9]. The proposed model can simultaneously predict flight delay, operational revenue, and carbon emissions, effectively explore the complex coupling relationships among these key indicators, and provide reliable data support and decision-making references for airline operation and management. This research contributes to promoting the digital transformation, intelligent operation, and green sustainable development of the civil aviation industry[10].Although the proposed model achieves good performance, there are still some limitations. For example, the dataset does not include weather, air traffic control, and airport operation factors. In the future, we will introduce multi-source spatiotemporal data such as weather radar and flow control to further improve prediction accuracy and build a more complete intelligent decision-making system for civil aviation.

References

[1] Bisandu B D ,Chandrakumar M ,Moulitsas I .Deep liquid neural network for prediction of weather-impacted flight delay[J].Intelligent Systems with Applications,2026,30200656.
[2] Shao L ,Huang Z ,He J , et al.A multi-factor flight delay prediction and evaluation method based on artificial neural network and cloud model[J].Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering,2026,240(5):860-870.
[3] Murad A ,Ruocco M .Pre-tactical flight-delay and turnaround forecasting with synthetic aviation data[J].CEAS Aeronautical Journal,2026,(prepublish):1-22.
[4] Liu M ,Wang J ,Lyu Z .A flight delay prediction model featuring a multi-strategy enhanced golf optimization algorithm integrated with LightGBM[J].Cluster Computing,2025,28(12):782.
[5] Yuan Y ,Wang Y ,Lai S C .Multi-Attribute Data-Driven Flight Departure Delay Prediction for Airport System Using Deep Learning Method[J].Aerospace,2025,12(3):246.
[6] Zhong Q ,Yu Y ,Huang Y , et al.Prediction and Optimization of Civil Aviation Flight Delays Based on Machine Learning Algorithms[J].International Journal of Computational Intelligence Systems, 2025,18(1):189.
[7] Yin C ,Du X ,Duan J , et al.Unveiling Hidden Dynamics in Air Traffic Networks: An Additional-Symmetry-Inspired Framework for Flight Delay Prediction[J].Mathematics,2025,13(14): 2274.
[8] Keyvandarian A ,Pirnia M ,Scott D .Robust optimal design of sustainable aviation fuel supply chain [J].Computers & Industrial Engineering,2026,216111929.
[9] Christian M ,Ireti A ,Christoph M .Online flight booking: digital nudging to decrease aviation-related carbon emissions[J].Information Technology & People,2024,37(1):29-50.
[10] Wang Y ,Sun H ,Lin Y , et al.An interval-valued estimation method of aircraft route carbon emission: A function of aircraft seat capacity and route flight time[J].Energy,2024,294130937.