Enhancing Aircraft Fuel Prediction with LSTM-Attention: Examining Lag Effects Across the Entire Flight

Binbin Ji[1*]**, Luqi Xu**[2]

[1] Department Information Management and Information Systems, Shanghai University, Shanghai 200444, Binbin_Ji@outlook.com
[2] Department Information Management and Information Systems, Shanghai University, Shanghai 200444, xu_qqq@outlook.com
*Corresponding author email: Binbin_Ji@outlook.com

Abstract: This research is centered on achieving full-flight fuel consumption fitting using a single model, with a specific focus on emphasizing the potential lag effects in aircraft fuel consumption. The proposed LSTM-Attention model integrates the capabilities of the LSTM network to effectively extract correlation features and sequence features from the data. Simultaneously, the attention mechanism assumes a crucial role in accentuating temporal dependencies and lag effects associated with fuel consumption in distinct flight segments. The parameters of the model undergo meticulous optimization through adjustment experiments. Experimental results demonstrate that compared to traditional models like BPNN, ELMAN, and RNN, the proposed model more efficiently extracts fuel consumption features throughout the entire flight, reducing the average prediction error by 66.5% and improving stability by an average of 38.8%. This model not only contributes to advancing fuel-saving research based on Quick Access Recorder (QAR) data but also holds promise for fault diagnosis applications in aviation.

Keywords QAR data; LSTM neural network; Attention mechanism; Full flight segment fuel consumption prediction

1. Introduction

The rapid growth of the aviation transportation industry, coupled with escalating international fuel prices and the continuous introduction of environmental policies, has prompted airlines to prioritize the control of fuel consumption as a primary task within their operational framework [1]. From an economic standpoint, fuel prices increased by 31.18% in 2019 compared to 2015 [2]. Globally, fuel costs, apart from labor expenses, constitute the highest proportion among various operational costs, averaging 23.2% of total expenses [3]. In the environmental context, fuel consumption directly contributes to the rise in carbon dioxide and greenhouse gas emissions. Consequently, the International Air Transport Association (IATA) [4] emphasizes the goal to reduce by 50% in carbon dioxide emissions by 2050 compared to 2005 levels. Growing global recognition of environmental conservation by international bodies, coupled with escalating fuel costs and ongoing demand for air transport, underscores the imperative for the effective implementation of fuel management strategies to ensure long-term sustainability.

While throttle optimization control technologies in flight management are maturing, solely relying on engineering innovations to enhance aircraft performance for reducing fuel consumption is becoming progressively challenging, with relatively high associated costs. Currently, both domestic and international airlines focus their fuel consumption reduction goals on accurately estimating aircraft fuel consumption and researching fuel-efficient flight operations to avoid the phenomenon of "burning fuel to carry fuel," thereby minimizing excess fuel consumption and enhancing fuel utilization efficiency. Neural networks, with their ability to deeply explore relationships within massive datasets, have become a focal point of research for estimating aviation fuel consumption, a complex nonlinear multivariate time series fitting problem. Trani et al. [5] first proposed using a backpropagation (BP) neural network to fit fuel flow for the ascent, cruise, and descent phases of flight in 2004. They employed parameters such as Mach number, weight, temperature, and altitude to construct the model, determined the optimal network topology through trial and error, and demonstrated that a three-layer BPNN trained with the Levenberg-Marquardt (LM) algorithm exhibited optimal performance. However, this method suffers from slow convergence and susceptibility to local optima. Building on this, Baklacioglu [6] improved Trani's work in 2016, using genetic algorithm-enhanced BP neural networks (GA-BP) with altitude and true airspeed as input parameters to determine the network's initial parameters, and separately fitting the ascent, cruise, and descent stages. Subsequently, Baklacioglu [7][8] in 2020 and 2021 respectively proposed the use of the CSA and PSO algorithms for optimization in the ascent and descent stages of the aforementioned model. The models exhibited unstable performance in test groups, and a complete enhancement in performance was not achieved.

These methods, however, neglect the time series characteristics of aviation data, failing to account for the temporal dependencies present in fuel consumption variations over time within each flight. In recent years, scholars both domestically and internationally have proposed various solutions for fuel consumption fitting based on the time and space characteristics inherent in Quick Access Recorder (QAR) data. Chen et al. [9] suggested using Long Short-Term Memory (LSTM) models to fit aircraft flow during the cruise phase, demonstrating superior performance compared to linear regression and random forest models, exhibiting high accuracy and generalization capability. Chen Cong et al. [10] proposed using a Gated Recurrent Unit (GRU)-improved Recurrent Neural Network (RNN) for fitting the entire flight phase. The improved GRU network achieved significantly higher prediction accuracy and curve-fitting capabilities. However, this experiment revealed that prediction deviations during the ascent and descent stages were noticeably greater than during the cruise phase. Elman neural networks have also been utilized for fuel prediction, with Partial Least Squares (PLS) regression and Particle Swarm Optimization (PSO) combined with the Elman neural network. This resulted in a PLS-PSO-Elman algorithm, reducing data dimensionality through PLS, simplifying network structure, and using PSO to optimize neural network connection weights, thresholds, and the number of hidden layer neurons. This addressed the deficiencies of incomplete training and poor learning accuracy in the Elman algorithm [11]. However, experimental results showed that although the new algorithm exhibited high computational efficiency and prediction accuracy, it still suffered from large variance in results, indicating instability in the model's predictions.

Upon analyzing and replicating the existing literature, three primary issues emerge:

(1) Some models overlook a certain degree of lag of aircraft fuel consumption typically exhibits. Actually, the fuel consumption rate of an aircraft typically exhibits a certain inertia, causing the fuel consumption not to immediately stabilize to a new level after adjusting thrust or speed [12]. Specifically, the fuel consumption of an aircraft is associated with various factors, including the flight phase, aircraft performance, environmental conditions, and more. Changes in these factors may not immediately manifest in the fuel consumption but involve a certain delay. For instance, during the aircraft's acceleration or climb phase, the engines may require more fuel to provide additional thrust, but the actual fuel consumption might peak after a certain period. Similarly, during the deceleration or descent phase, reducing thrust may not immediately result in a corresponding decrease in fuel consumption, exhibiting a certain response lag.

(2) Some studies often narrow their focus to fitting fuel consumption models for specific flight segments, neglecting the need for a unified model that covers the entire flight duration. This segmented approach results in operational inefficiencies and impracticality for real-world applications.

(3) Some models that focus on the full-flight fitting cannot effectively extract distinct features for different flight segments, leading to issues where the prediction accuracy of one or several segments is significantly lower than the rest, limiting their overall effectiveness.

To address these limitations, we propose the LSTM-Attention model. The innovation of this model is that it is an integrated model focusing on entire flight segment fuel fit with attention to segment differences, taking into account the delay phenomenon in aircraft fuel consumption. The LSTM model is employed to extract both time-series features and inter-parameter correlations. Furthermore, the attention mechanism is utilized to emphasize and identify distinct temporal dependencies and sequence delays in fuel consumption across different flight segments. This model achieves a unified approach to predict aircraft fuel consumption throughout the entire flight phase while capturing variations in segment characteristics.

The remaining sections of this paper are organized as follows. The second section explains the structure and methodology of model establishment. The third section discusses the experimental process and results. Ane the fourth section summarizes our study.

2. Fuel Consumption Model

The LSTM-Attention model proposed in this paper first inputs the constructed QAR time series data into the LSTM network to extract temporal features and the interrelated features among various parameters. Subsequently, attention mechanism operations are performed on the hidden layer outputs at each time step. This mechanism enables the model to focus on and identify the temporal dependencies and delays in fuel consumption during different flight phases. The specific model structure is illustrated in Figure 1.

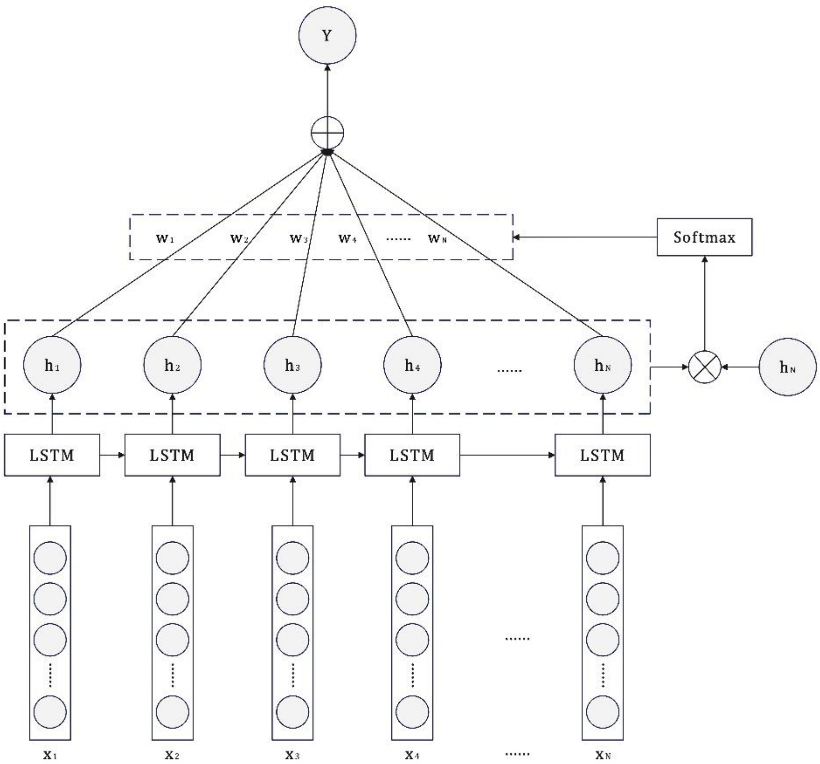

Figure 1. LSTM-Attention Structure

2.1. LSTM

LSTM (Long Short-Term Memory) is a type of recurrent neural network that is capable of processing sequential data with relatively long intervals and delays. The neural unit of LSTM is composed of memory cell and three gates, including input gate, forget gate and output gate [13]. The memory cell is set for preserving important information; the input gate determines whether the new information should be stored into the memory cell; the forget gate decides whether to forget the information in the memory cell; and the output gate decides whether to use the information in the memory cell as the current output. The management of these gates can effectively capture important long-term dependencies in the sequence and can solve the gradient problem. The following formulas [13] represent the logic of each LSTM unit. Set the input time series be x and t be the current moment; $\tilde{C}_t$ indicates the candidate cell state; C_t indicates the cell state; i_t、 o_t and f_t indicate the input gate, output gate, and forget gate respectively.

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (1)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tanh(W_C[h_{t-1}, x_t] + b_C) \quad (3)$$

$$C_t = i_t \cdot \tilde{C}_t + f_t \cdot C_{t-1} \quad (4)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (6)$$

Here, σ represents sigmoid activation function, W and b represent the weight and bias.

2.2. Attention Mechanism

The attention mechanism, mimicking the attention allocation mechanism of biological observation behavior, perform the importance evaluation of information in the sequence through probability calculation to highlight important one and selectively reduce or even ignore unimportant one [14]. See the following formula for specific Equation (7)：

$$Attention(Q, K, V) = Softmax(Q \times K^T) \cdot V \quad (7)$$

In the calculation, Q stands for query, K for key, and V for value. In this paper, Q represents the hidden layer state of the last time step of LSTM, K represents the hidden layer state of all time steps of LSTM, and V is the same as K.

The essence of the LSTM-Attention model lies in integrating the Attention mechanism into the LSTM network, enabling the network to autonomously grasp and prioritize crucial information. Within the LSTM-Attention model, a weight vector is computed for the output of each time step using the attention mechanism. This weight vector signifies the significance of each element in the input sequence at the current time step.

3. Experimental work

3.1. Dataset

This study focuses on the research of aircraft fuel consumption rates, utilizing data from the Quick Access Recorder (QAR) for experimentation. Based on the principles of flight control [15], 17 parameters related to flight status, engine performance, and environmental factors were selected from the data for experimentation. The specific parameters and their meanings are detailed in Table 1. Among them, the sum of the fuel flow rates of Engine 1 (FF1C) and Engine 2 (FF2C) is the predictive parameter, while the remaining parameters are state parameters.

Table 1 Model Parameters

Parameter Dimension	**Parameter Name**	**Parameter Annotations**	**Parameter Units**
Flight Status	ALT_STDC	Altitude Standard	FEET
	TAS	True Airspeed	KNOT
	PITCH	Pitch Angle	G
	VRTG	Normal Acceleration	G
	LONG	Longitudinal Acceleration	G
	GWC	Gross Weight in Ton (1000 Kg)	t
Engine Performance	N11	CFM N1 Actual (Low Rotor Speed) ENG 1	% RPM
	N21	CFM N2 Actual (High Rotor Speed) ENG 1	% RPM
	CK_EGT1	Cockpit - EGT Engine 1	deg C
	N12	CFM N1 Actual (Low Rotor Speed) ENG 2	% RPM
	N22	CFM N2 Actual (Low Rotor Speed) ENG 2	% RPM

	CK_EGT2	Cockpit - EGT Engine 2	deg C
Environment Factors	WIN_SPDR	Wind Speed	KTS
	WIN_DIR	Wind Direction	DEG
	SAT	Outside Static Temperature	DEG
Fuel Consumption Rate	FF1C	Fuel Flow Engine 1	Kg/h
	FF2C	Fuel Flow Engine 2	Kg/h

To mitigate the impact of differing data scales on the experimental outcomes, the present study employed the min-max scaling method to normalizing each parameter. This normalization process scales the data to [0,1]. And the calculation process is expressed by Equation (8).

$$x_i' = \frac{x_i - \min(x)}{\max(x) - \min(x)} \tag{8}$$

Here, x_i denotes the original data, and x_i' represents the normalized data, $\min(x)$ and $\max(x)$ respectively denote the maximum and minimum values of a specific parameter.

3.2. Wavelet Domain Denoising

Using actual data from the Quick Access Recorder (QAR) can alleviate the issue of model inefficiency caused by inaccurate simulated flight data. However, QAR data, being sensor data, may contain outliers or noise. *Wavelet domain denoising* involves decomposing the signal into high-frequency and low-frequency components based on wavelet basis functions, filtering noise by selecting an appropriate threshold, and reconstructing the signal after multi-level decomposition [17]. In this study, the Haar wavelet basis function was chosen for a 5-level decomposition, and a hard threshold function was employed for signal filtering. Figure 2 illustrates the before-and-after comparison of the wind speed (WIN_SPDR) after denoising. The results demonstrate that wavelet domain denoising effectively eliminates peaks while ensuring the smoothness and similarity of data.

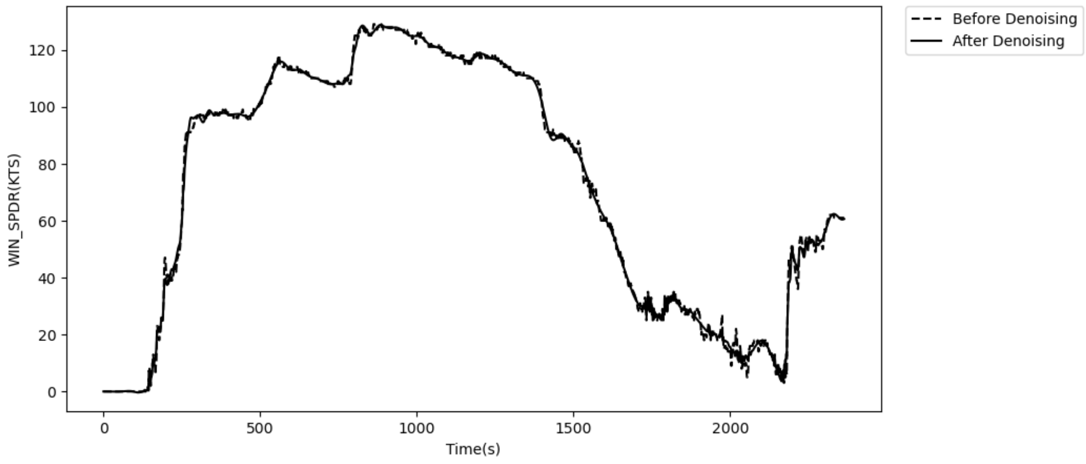

Figure 2. Denoising Comparison: Before and After

3.3. Model Parameter Selection and Training

3.3.1. Hyperparameter Selection

The network relies on gradient descent for weight updates. Learning rate is one of the most critical hyperparameters which determines the convergence of model. An excessively large learning rate may lead to difficulties in convergence, while a too-small learning rate may result in slow convergence. Moreover, using a fixed learning rate throughout the training process can lead to issues of large adjustment magnitudes later in the training. Therefore, this study employs the Adam algorithm to optimize network parameters. The Adam algorithm adjusts the learning rate based on the state of the training loss function that is suitable for model involving large-scale data and parameters [18].

In addition, key parameters affecting the training effectiveness of the model also include batch size, the number of hidden layer nodes, and the length of the time series. A larger batch size leads to a descent direction closer to the gradient descent direction, resulting in faster convergence. However, higher memory utilization is incurred. Therefore, this study sets the batch size to 256. And other crucial parameters are determined through parameter tuning experiments.

To determine the optimal values for the learning rate and time series length, adjustment experiments are conducted with a fixed number of hidden layer nodes. Each set of parameters is input into the network for 10 computations, and the average loss function value is taken as the experimental result for that set of parameters. Figure 3 displays the outcomes of the tuning experiments. The findings reveal that the validation average MSE reaches its minimum when the learning rate is set to 1E-4 and the time series length is 10.

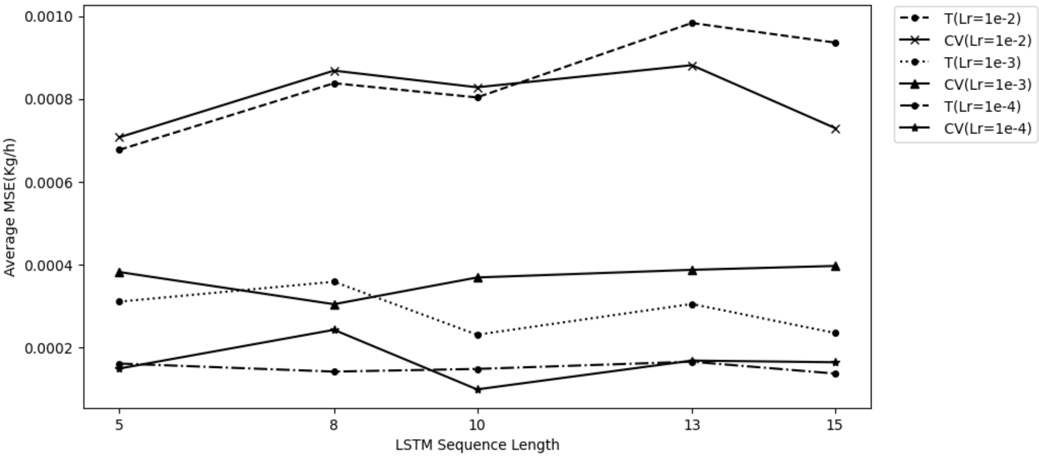

Figure 3. Training and Cross-validation Average MSE Results across Various Learning Rates and Sequence Lengths

Next, with the learning rate and time series length fixed at the previously determined optimal values from the initial experiment, an experimentation is performed to ascertain the optimal number of hidden layer nodes. Adhering to the empirical formula (9), the range for the number of hidden layer nodes is established as [5, 14]. Similarly, the average loss MSE value from 10 computations for each set of parameters is taken as the experimental result. The results of this experiment are presented in Figure 4.

$$hidden\_size = \sqrt{n+m} + a \tag{9}$$

Here, n represents the number of input parameters, m represents the number of output parameters, and a signifies the adjustable parameter with a range of [1, 10].

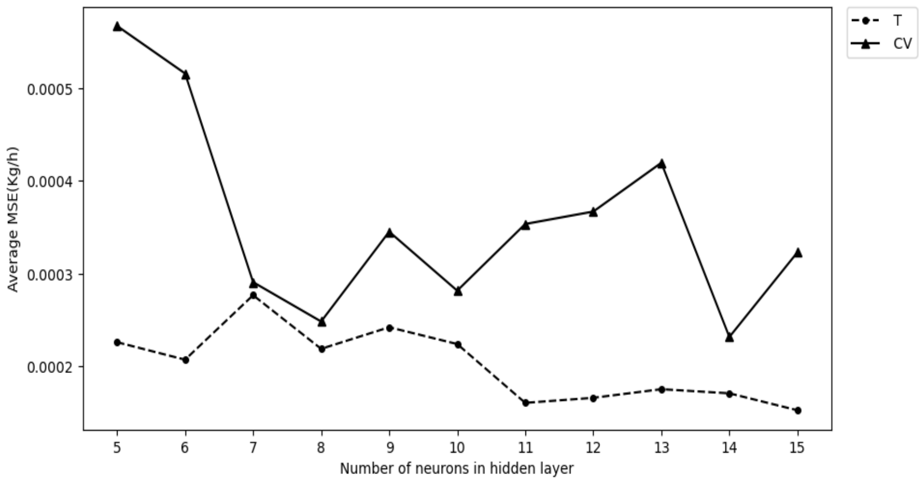

Figure 4. Training and Cross-validation Average MSE Results Across Various NN Topologies

The results indicate that the minimum validation MSE value is achieved when the number of hidden layer nodes is set to be 14, making it the optimal choice.

3.3.2. Aircraft Fuel Consumption Prediction Curve

To validate the accuracy and utility of the model, this study conducted a comparative analysis between actual fuel consumption values and predicted values using real flight data. Figure 5 presents the results of fuel consumption reconstruction for the test data. A congruence between the trends of actual and predicted fuel consumption, with the model accurately capturing real-time variations, validates the accuracy and applicability of the model.

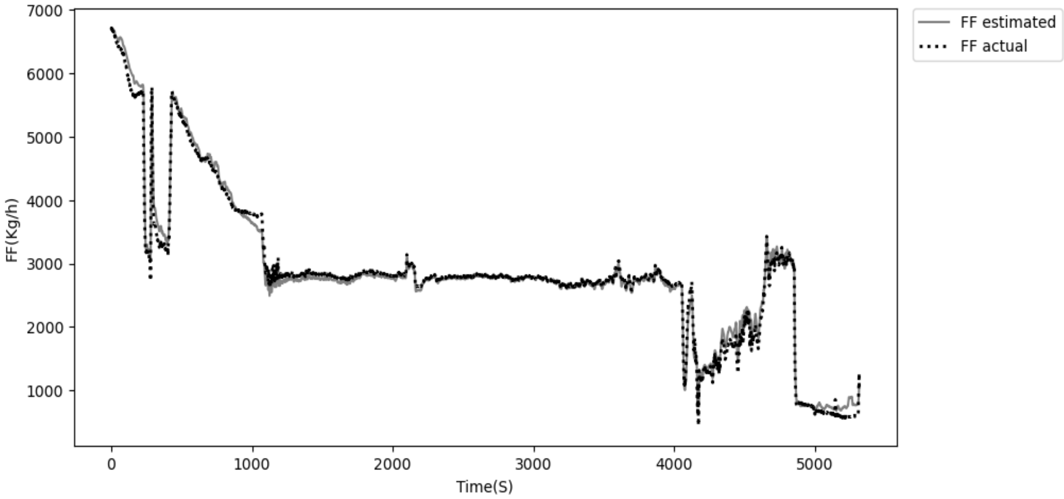

Figure 5. The Comparison of Estimated and Actual Fuel Flow Rates Values

3.4 Model Validation and Analysis

In order to quantitatively analyze the performance of the model, this study employs the mean and standard deviation of the Mean Squared Error (MSE), as well as the Root Mean Squared Error (RMSE) as evaluation metrics [19]. The calculation formulas are provided in Equations (10) to (13). The mean of MSE, as well as a smaller standard deviation of MSE, indicates a more accurate and stable model.

$$MSE = \frac{1}{n} \times \sum_{i=1}^{n} (Y_i - Y_i')^2 \tag{10}$$

$$SE\_Mean = \frac{1}{m} \times \sum_{i=1}^{m} MSE_i \tag{11}$$

$$MSE\_StDev = (\frac{1}{m} \times \sum_{i=1}^{m} (MSE_i - MSE\_Mean)^2)^{1/2} \tag{12}$$

$$RMSE = (\frac{1}{n} \times \sum_{i=1}^{n} (Y_i - Y_i')^2)^{1/2} \tag{13}$$

Here, n represents the sample size, m denotes the number of calculations, Y_i represents the actual value, and Y_i' represents the predicted value.

The error metrics for this model and five comparison models are calculated as shown in Table 2. The results indicate that, compared to the BP neural network, the ELMAN, RNN, and GRU networks, which focus on data sequence features, exhibit an average reduction in prediction error of approximately 20.6%. However, their model stability decreases by approximately 56.4%. The LSTM network, in comparison, achieves a 59.9% reduction in prediction error and an 84.2% improvement in stability compared to the aforementioned three networks. The LSTM model with the integrated attention mechanism demonstrates the highest accuracy and stability. Compared to the LSTM, the prediction error decreases by 37.7%, and stability improves by 2%. Compared to all five models, the average reduction in prediction error is 66.5%, and the average improvement in stability is 38.8%.

Table 2 Error metrics of the model and five comparison models

	MSE		RMSE
	Mean	**StDev**	
GABP	5.19E-04	9.99E-05	2.41E-02
ELMAN	4.14E-04	1.87E-04	2.03E-02
RNN	3.43E-04	1.16E-04	1.85E-02
GRU	4.48E-04	2.28E-04	2.12E-02
LSTM	1.57E-04	2.50E-05	1.25E-02
LSTM-Attention	**9.77E-05**	**2.45E-05**	**9.88E-03**

4. Conclusion and further suggestion

The aircraft fuel consumption exhibits a lag effect, and the lag effects vary across different flight segments. Additionally, flight data possesses both spatial and temporal characteristics. In this paper, we propose an LSTM-Attention neural network capable of effectively extracting data sequence features. The model employs an attention mechanism to adjust the focus on the temporal dependencies and lag effects in different flight segments. Prediction analysis on real flights demonstrates that compared to models like BPNN and ELMAN, our proposed model can more effectively extract the correlation and sequence features within QAR data. It achieves high-precision and high-stability full-flight fuel flow fitting, showcasing excellent generalization and robustness.

The model not only exhibits exceptional performance in experimental scenarios but also holds substantial promise for practical applications. It can be applied to real-world aviation challenges such as fuel efficiency studies and fault diagnosis, offering a viable tool for airlines to improve fuel efficiency in operational flights. Future research directions include further optimizing the model architecture, incorporating additional flight data features, and conducting large-scale validation experiments in actual aviation operations to validate the reliability and practical effectiveness of the model.

In summary, this study presents a novel and effective modeling approach for predicting aviation fuel flow rates, contributing significantly to the achievement of energy-saving and environmentally friendly goals in the aviation industry.

Acknowledgement

Binbin Ji and Luqi Xu contributed equally to this work and should be considered co-first authors.

References

[1] Khan W. A., et al., “A Novel Self-Organizing Constructive Neural Network for Estimating Aircraft Trip Fuel Consumption” [J]. *Transportation Research Part E: Logistics and Transportation Review*, 2019, 132(01): 72–96.

[2] IATA, Fuel efficiency [EB/OL]. https://www.iata.org/en/programs/ops-infra/fuel/fuel-efficiency/

[3] IATA, IATA Forecast Predicts 8.2 billion Air Travelers in 2037 [EB/OL]. https://www.iata.org/en/pressroom/pr/2018-10-24-02/

[4] Trani A., et al., “A neural network model to estimate aircraft fuel consumption” [J]. *AIAA 4th Aviation Technology, Integration and Operations (ATIO) Forum*, 2004: 6401.

[5] Baklacioglu T., “Modeling the Fuel Flow-Rate of Transport Aircraft during Flight Phases Using Genetic Algorithm-Optimized Neural Networks” [J]. *Aerospace Science and Technology*, 2016, 49(1): 52–62.

[6] Oruc R., et al., “Modelling of Fuel Flow-Rate of Commercial Aircraft for the Climbing Flight Using Cuckoo Search Algorithm” [J]. *Aircraft Engineering and Aerospace Technology*, 2020, 92(3): 495–501.

[7] Oruc R., et al., “Modeling of Fuel Flow-Rate of Commercial Aircraft for the Descent Flight Using Particle Swarm Optimization” [J]. *Aircraft Engineering and Aerospace Technology*, 2021, 93(2): 319–326.

[8] Luo W. Z., et al., “An Aircraft Fuel Flow Model of Cruise Phase Based on LSTM and QAR Data” [J]. *2020 13th International, Symposium on Computational Intelligence and Design (ISCID)*, 2020(13): 12-13.
[9] Chen C., et al., “Aircraft Fuel Flow Prediction Based on Improved GRU Neural Network” [J]. *Science Technology and Engineering*, 2021, 21(27): 11663-11673.
[10] Chen C., et al., “Aircraft Fuel Flow Prediction Based on PLS-PSO-Elman” [J]. *Science Technology and Engineering*, 2021, 21(05): 1963-1973.
[11] Jia L. Y, et al. “Fine-grained modeling and validation of transient states in turbofan engines” [J]. *Propulsion Technology*, 2023, 44(05): 48-56.
[12] Hochreiter S., & Schmidhuber J., “Long Short-Term Memory” [J]. *Journal of Noural Computation*, 1997, 9(8): 1735-1780.
[13] Shi L., Wang Y., Cheng Y., & Wei R. B., “A Review of Attention Mechanism Research in Natural Language Processing” [J]. *Data Analysis and Knowledge Discovery*, 2020, 4(05), 1-14.
[14] Yang H., & Wang L. J., “QAR data outlier detection algorithm based on clustering and fitting” [J]. *Computer Engineering and Design*, 2015, 36(01): 174-177.
[15] Yang H., & Wang L. J., “QAR data outlier detection algorithm based on clustering and fitting” [J]. *Computer Engineering and Design*, 2015, 36(01): 174-177.
[16] Chen C., et al., “Aircraft Fuel Flow Prediction Based on QAR Data” [J]. *Control Engineering*, 2019, 26(04): 752-758.
[17] Liu R., Liu Y., & G, J. G., “Improved AdaNet Based on Adaptive Learning Rate Optimization” [J]. *Computer Applications*, 2020, 40(10): 2804-2810.
[18] Khan W. A., et al., “Prediction of Aircraft Trajectory and the Associated Fuel Consumption Using Covariance Bidirectional Extreme Learning Machines” [J]. *Transportation Research Part E: Logistics and Transportation* Review, 2021,145(01):102-189.