Fuel Consumption Model of Aircraft in Descent Stage Based on DBN

Zhentao Wu*, Xueren Li, Jun Du
Aeronautics Engineering College, Air Force Engineering University, Xi’an 710038, China
*Corresponding author’s e-mail: shus662@qq.com

Abstract. Due to the pressures of the current ecological environment and the rise of fuel prices, it is necessary to calculate the fuel volume of aircraft accurately. In order to calculate the fuel consumption of a flight, the key is to establish an accurate fuel consumption model. In the process of descent, because the environment around the aircraft changes dramatically, compared with other stages, the fuel consumption factors affecting the descent stage will be more. But at present, the domestic fuel consumption model for aircraft descent stage is not accurate enough. To solve this problem, a method of building fuel consumption model in aircraft descent stage based on flight data using Deep Belief Network (DBN) is proposed. Firstly, the parameters related to fuel consumption are selected from the flight data, then the correlation between each parameter and fuel consumption rate is calculated by mutual information algorithm, and finally the parameters with the highest correlation are selected for model training. Compared with the traditional BP neural network and Echo State Network (ESN), the accuracy has been greatly improved.

1. Introduction
With the development of the society, people are becoming more and more aware of energy conservation and emission reduction, eager to have a better living environment, and most of the environmental pollution is due to exhaust emissions. Aircraft as a large transport tool, its pollution emissions cannot be underestimated. Taking Guangzhou Baiyun Airport as an example, the daily emissions generated by aircraft are equivalent to the emissions of 600,000 taxis. Carbon emissions from civil air transport account for about 2% of the total carbon emissions from human activities [1]. Therefore, the problem of energy saving and emission reduction needs to be solved urgently. At present, the refueling of aircraft has become one of the core tasks. How to control the refueling quantity of aircraft has become a difficult problem faced by major airlines. For military aircraft, it is also very important to realize accurate refueling. If the fuel is too much, it will affect the maneuverability of the fighter in the execution of combat operations, and impact the structure of the fighter itself when landing, thus reducing the life of aircraft; if the fuel there is too little, it will not be able to complete the tactical requirements, even lead to no return or the crash of the aircraft. Therefore, in order to achieve accurate refueling of aircraft, it is necessary to establish an accurate fuel consumption model.

There are several methods to build the fuel consumption model of aircraft:

(1) Modeling based on aircraft performance. Collins [2] put forward the principle of energy conservation, and established the fuel consumption model of the aircraft by analyzing the changes of kinetic energy and potential energy of the aircraft. Wang [3] proposed to use the energy state method and optimize it through Genetic Algorithm (GA), but all of these methods need to get the performance data of the aircraft by consulting the technical manuals, which is more cumbersome and less accurate.
(2) The method based on multiple regression. Tian [4] proposed the establishment of fuel consumption model for non-standard actions of military aircraft by means of multiple linear regression. But for fuel consumption model, it is non-linear. Simple linear regression method cannot achieve better results. T. Baklacioglu [5] innovatively proposed the use of Genetic Algorithm (GA), and obtained the expression between fuel consumption rate and height and true air speed in climbing stage through parameter identification.

(3) BADA-based method. BADA is a database of basic performance parameters of aircraft developed by EUROCONTROL. It contains most of the civil aviation aircraft nowadays. Wu [6] uses BADA model to simulate and realize the calculation of pollutant emission during aircraft descent.

(4) The method based on Neural Network. Trani A [7] proposed the use of Artificial Neural Network (ANN) algorithm to calculate fuel consumption in 1997. But its convergence speed is slow and it is easy to fall into local optimum; Liu [8] uses BP Neural Network based on QAR data to model the whole range of the aircraft; Wang [9] uses Echo State Network (ESN) method to calculate the fuel consumption of the whole range based on trajectory data. However, the above methods have great errors in establishing the fuel consumption model of aircraft in the descent stage.

In 2006, Hinton first put forward the idea of Deep Learning (DL) in Science, which attracted the attention of many experts and scholars. Deep Learning can achieve the approximation of complex functions by learning a deep nonlinear network structure. For fuel consumption model, there are many influencing factors, especially in the descent stage. The drastic changes of the surrounding environment will increase the complexity of the model. Therefore, based on flight data, a fuel consumption model of a certain aircraft in descent stage is established by using Mutual Information (MI) to select relevant parameters and Deep Belief Network (DBN) in this paper.

2. Correlation model of fuel consumption during aircraft descent

At present, Pearson correlation coefficient method, grey correlation method [10] and mutual information method [11] are widely used in the calculation of correlation degree. However, Pearson coefficient is only sensitive to the trend between the two sets of data, but not sensitive to the absolute value, while the theoretical research of grey correlation method is not deep enough, and the subjectivity is too strong to achieve satisfactory results.

Mutual Information (MI) is a method to measure the degree of correlation between a set of samples on two attributes. It is often used in feature selection and decision tree problems. Unlike correlation coefficient, it is not limited to real random variables, but reflects the similarity of joint distribution \( p(X,Y) \) and edge distribution \( p(X) p(Y) \). It came into being after Shannon [12] put forward the concept of entropy, which was used to measure the correlation between two random variables. The information entropy \( H(X) \) of a discrete random variable \( X \) can be defined as:

\[
H(X) = -\sum_{x \in X} p(x) \log p(x)
\]  
(1)

where \( p(x) \) represents probability distribution function of \( X \).

Generally, mutual information can be defined as:

\[
I(X;Y) = \sum_{x,y} p(x,y) \log \left( \frac{p(x,y)}{p(x)p(y)} \right)
\]  
(2)

where \( p(x,y) \) is the joint probability distribution function of \( X \) and \( Y \), \( p(x) \) and \( p(y) \) is the edge probability distribution function of \( X \) and \( Y \).

The greater the mutual information value \( I(X;Y) \) between the two variables, the stronger correlation between them.

3. Brief Introduction of Deep Belief Network
DBN is stacked by many Restricted Boltzman Machines (RBM). Through continuous training and updating of weights at all levels, the error of the whole network is minimized.

3.1. Restricted Boltzman Machines
Constrained Boltzmann Machine is proposed by Hinton [13]. It is a Generative Stochastic Network (GSN). Its network structure is shown in Figure 1.

![Network structure of RBM.](image)

As can be seen intuitively from Figure 1, RBM consists of two layers, one is the visual layer and the other is the hidden layer. There is no connection between layers, but there is full connection between layers, and the corresponding value of each neuron can only be 0 or 1, 0 means that the neuron is not activated, 1 means that the neuron has been activated. Assuming that there are $n$ neurons in the visual layer $v$ and $m$ neurons in the hidden layer $h$, the energy function can be defined as:

$$E(v, h) = - \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} v_i h_j - \sum_{j=1}^{m} c_j h_j - \sum_{i=1}^{n} v_i b_i$$

(3)

where $v_i$ is the $i$-th neuron of the visual layer, $h_j$ the $j$-th neuron of the hidden layer, $b_i$ is the bias of the $i$-th visual layer, $c_j$ is the bias of the $j$-th hidden layer, $w_{ij}$ is the weight matrix of the connection between the visual layer and the hidden layer.

The energy function $E(v, h)$ determines the error in the process of the conversion between the specific features and the abstract features. The larger the value of energy function, the more chaotic or dispersed the system is; the smaller the value of energy function, the more orderly or concentrated the system is. Then the edge probability distribution of hidden layer nodes and visual layer nodes can be obtained by calculating the energy function of the system:

$$P(v_i | h) = \frac{e^{-E(v_i, h_j, \theta)}}{\sum_h e^{-E(v_i, h_j, \theta)}}$$

(4)

$$P(h_j | v_i) = \frac{e^{-E(v_i, h_j, \theta)}}{\sum_h e^{-E(v_i, h_j, \theta)}}$$

(5)

then the joint probability distribution of the hidden layer and the visible layer is:

$$P(v, h) = \frac{e^{-E(v, h, \theta)}}{Z(\theta)}$$

(6)

where $Z(\theta)$ is normalization factor, its concrete expression is as follows:

$$Z(\theta) = \sum_{v, h} e^{-E(v, h, \theta)}$$

(7)
The corresponding conditional probability distribution function can be obtained by calculating the state of neurons and \( \theta = \{ h, v, w, b, c \} \).

\[
P(h_j = 1|v) = \text{sigmoid}\left( \sum_{i=1}^n w_{ij} v_i + b_i \right) \quad (8)
\]

\[
P(v_i = 1|h) = \text{sigmoid}\left( \sum_{i=1}^n w_{ij} h_j + c_j \right) \quad (9)
\]

where \( \text{sigmoid} = \frac{1}{1 + e^{-\eta}} \) is an activation function to activate neurons. According to the above, the edge probability values of neurons can be obtained:

\[
P(v; \theta) = \frac{\sum_v e^{-E(v, h; \theta)}}{\sum_v \sum_h e^{-E(v, h; \theta)}} \quad (10)
\]

\[
P(h; \theta) = \frac{\sum_h e^{-E(v, h; \theta)}}{\sum_v \sum_h e^{-E(v, h; \theta)}} \quad (11)
\]

Due to the structural characteristics of RBM itself, the interlayer neurons are fully connected, while the intralayer neurons are not connected, so that the complexity of RBM stacking does not increase dramatically, only through training among multiple RBM layers.

3.2. Deep Belief Network

DBN is a classical Deep Learning method. It uses multi-layer RBM stacking and training layer-by-layer. It obtains the joint probability distribution of training samples, and then estimates the conditional probability distribution. The network structure of DBN is shown in Figure 2.

![Figure 2. Network structure of DBN.](image)

As can be seen from the above, the DBN model consists of the input layer at the bottom, the hidden layer at the middle and the output layer at the top. Because of the stacking of RBM layers, the hidden layer neurons of the former RBM become the visual layer neurons of the latter RBM. The final output layer is a layer of BP neural network. The RBM neurons of the last layer are used as the initial value of BP neural network training, and the whole network model is fine-tuned by supervised learning. That is to say, the error is propagated from the top to the bottom by using the back propagation characteristics of BP itself.

In general, the training process of DBN is divided into two stages: Pre-training stage and Fine-training stage. Pre-training stage is an unsupervised network training, while fine-training stage is a supervised network training. In the pre-training stage, through the propagation between layers of RBM, the final propagation to the top layer becomes the initial value of BP neural network. In the fine-training
stage, the training error is propagated back through BP neural network, which realizes the fine-training of the whole network model structure, so that the final network structure can be more accurate.

The ultimate training objective of DBN model is to solve the parameters \( \theta = \{h, v, w, b, c\} \). Generally, the maximum likelihood function \( L(\theta) \) is used to calculate the parameters.

\[
L(\theta) = -\log P(y; \theta | x)
\]

(12)

The estimation of the parameter \( \theta \) Obtained from the upper formula:

\[
\theta = \theta + \epsilon \cdot \frac{\partial \ln L}{\partial \theta}
\]

(13)

Therefore, by using the maximum likelihood function and its estimation, we can get the connection weights \( w_{ij} \) between the neurons in the visual layer and the hidden layer, as well as the bias of the neurons in the visual layer and the hidden layer, that is, the bias of the neurons in the visual layer \( b_i \) and the bias of the neurons in the hidden layer \( c_j \).

\[
\frac{\partial L(\theta)}{\partial w_{ij}} = \Delta w_{ij} = \langle v_i h_j \rangle_{\text{data}} - \langle v_i h_j \rangle_{\text{model}}
\]

(14)

\[
\frac{\partial L(\theta)}{\partial b_i} = \Delta b_i = \langle v_i h_j \rangle_{\text{data}} - \langle v_i h_j \rangle_{\text{model}}
\]

(15)

\[
\frac{\partial L(\theta)}{\partial c_j} = \Delta c_j = \langle v_i h_j \rangle_{\text{data}} - \langle v_i h_j \rangle_{\text{model}}
\]

(16)

where \( \langle \cdot \rangle_{\text{data}} \) is the expected values calculated from raw data, \( \langle \cdot \rangle_{\text{model}} \) is the expected values calculated from model.

4. Experiments and analysis

4.1. Data selection

In this paper, the flight parameters of 55 flights of an aircraft in one year are selected, of which 50 flights are used as training set and the remaining 5 flights are used as test set. The parameters included in the data are fuel consumption rate, pitch angle, longitudinal acceleration, normal acceleration, relative pressure height, exhaust temperature, true air speed, low-pressure rotor speed, high-pressure rotor speed and atmospheric static temperature. Among the selected parameters, except that the fuel consumption rate directly reflects the flow rate of fuel consumption, the other parameters have a certain relationship with the fuel consumption rate. In order to reduce the external factors caused by the weather, the selected training set data are distributed every month of the year. In addition, because the sensors used in the acquisition of various parameters are different, the sampling rate is also different, so it is necessary to unify the sampling rate, so that their dimensions are unified, and the final sampling time is 0.5s.

4.2. Experimental steps

The specific experimental steps of this experiment are as follows:

1. Determine input and output. Output is the fuel consumption rate, which reflects the flow of fuel consumption. There are 9 inputs, among which the high-pressure rotor speed and the low-pressure rotor speed are highly correlated, so one of them is removed to reduce the dimension of the input variables. So now the input is 8 parameters.

2. Establish the correlation model of fuel consumption in descending stage. The remaining 8 parameters are calculated with MI with fuel consumption rate respectively, and the final results are sorted out. Several parameters with the highest correlation are selected as the input of the follow-up training model.

3. DBN model training. The final selected parameters are put into the DBN model for training.
(4) Model validation. The test set data are input into the trained DBN model, and the estimated fuel consumption rate is obtained. Then the training effect of the model is verified by comparing with the real value.

4.3. Evaluation index of experiment
In order to test the accuracy of fuel consumption rate trained by DBN model, relative error $E$ and determinant coefficient $R^2$ are introduced to evaluate the network model. Specific formulas are shown in (18) and (19):

$$E = \frac{|\hat{y}_i - y_i|}{y_i}$$

$$R^2 = \frac{\left(\sum_{i=1}^{l} \hat{y}_i y_i - \sum_{i=1}^{l} \hat{y}_i \sum_{i=1}^{l} y_i\right)^2}{\left(\sum_{i=1}^{l} \hat{y}_i^2 - \left(\sum_{i=1}^{l} \hat{y}_i\right)^2\right) \left(\sum_{i=1}^{l} y_i^2 - \left(\sum_{i=1}^{l} y_i\right)^2\right)}$$

where $l$ is the sample number of test sets, $\hat{y}_i, i = 1,2,\ldots,n$ is the $i$-th predict value of fuel flow rate, $y_i, i = 1,2,\ldots,n$ is $i$-th real flight data, $n$ is the number of samples. The smaller the relative error $E$ is, the closer the determinant $R^2$ ($0 < R^2 < 1$) is to 1, the better fitting degree of DBN model is.

4.4. Experimental process

4.4.1. Relevance Model of Fuel Consumption in Downward Stage. The parameters of pitch angle, longitudinal acceleration, normal acceleration, relative pressure height, exhaust temperature, true air speed, high-pressure rotor speed and atmospheric static temperature are calculated with fuel consumption rate respectively. The results are shown in Figure 3.

![Figure 3](image)

Figure 3. MI between each parameter and fuel flow rate.
As can be seen from the above, the correlation between true air speed, exhaust temperature, high-pressure rotor speed, atmospheric static temperature and relative pressure height and fuel consumption rate is relatively high, so these 5 parameters are selected as the input of DBN to train the network model.
The purpose of MI calculation is to reduce the dimension of input. If the dimension of input vector is high, it will increase the complexity of the model, resulting in too long training time, and the parameters with low correlation will also affect the accuracy of the training model. Therefore, only a few parameters with high correlation are selected as the output of DBN.

4.4.2. Training of DBN model. Because MI calculation has been done before, true air speed, exhaust temperature, high-pressure rotor speed, atmospheric static temperature and relative pressure height are selected as the input of DBN to train the network model. Its symbolic representation and unit are shown in Table 1.

| Parameters                      | Symbolic Representation | Unit     |
|---------------------------------|-------------------------|----------|
| True air speed                  | TAS                     | km/h     |
| Exhaust temperature             | T4                      | °C       |
| High-pressure rotor speed       | N2                      |          |
| Atmospheric static temperature  | T                       | °C       |
| Relative pressure height        | H                       | m        |
| Fuel flow rate                  | ff                      | kg/h     |

Firstly, normalize the input and output data, and then initialize the parameters of DBN model. The number of RBM layers is 2. Because there are 5 inputs, so the number of input layer nodes is 5. Set 10 neurons in the first layer, the number of neurons in the second layer is 5, the number of nodes in the output layer is 1, the learning rate is 0.001, and the number of iterations is 100. After the training of the network model, the test data are normalized and input into the network model, and then output the corresponding results. Finally, the estimated fuel consumption rate is compared with the actual flight data. Taking one of the flights as an example, the simulation result is shown in Figure 4.

![Comparison of DBN prediction results with actual results.](image)

The prediction result of DBN shows that the accuracy of the model is very high. The relative errors $E$ and determination coefficients $R^2$ between the prediction results of DBN of five flights and the real records are shown in Table 2.

|     | $E$     | $R^2$  |
|-----|---------|--------|
| 1st | 0.0055822 | 0.9617 |
As can be seen from Table 2, the accuracy of the model of DBN is very high, and the $R^2$ can reach about 0.96. So it can be concluded that the DBN method is very effective for the establishment of fuel consumption model in aircraft descent stage.

### 4.4.3. Comparison with BP and ESN

In order to further highlight the superiority of DBN model, BP neural network and ESN are used to model and validate the same sortie aircraft. Because of the large number of sampling points in the descent stage, the comparison of data in the whole descent stage is not clear enough, so the prediction effect of 100 points is used to show, as shown in Figure 5.

![Figure 5. Comparison of DBN, ESN, BP.](image)

Fig. 5 shows the prediction results of the three methods. In the process of increasing fuel flow, ESN has the best predictive effect, while DBN has better predictive effect than the other two when it tends to stabilize gradually, and BP has the worst effect among the three. The specific prediction results are shown in Table 3.

| Predicted oil volume (kg) | Actual oil volume (kg) | $E$ | $R^2$ |
|--------------------------|------------------------|-----|-------|
| BP                       | 1489                   | 1628| 0.0853808 | 0.7221 |
| ESN                      | 1553                   | 1628| 0.0460688 | 0.8104 |
| DBN                      | 1618                   | 1628| 0.0061425 | 0.9576 |

Compared with ESN and BP, the prediction error of DBN algorithm is greatly reduced. Therefore, for the fuel consumption model of aircraft descent stage, the DBN method is more accurate. The shallow neural networks such as BP and ESN cannot achieve good results. Perhaps in the cruise phase, which is relatively stable, the effect may be good, but in the descent phase, when the environment changes dramatically, the use of DBN method has obvious advantages.

### 5. Conclusion

In order to solve the problem of inaccurate establishment of fuel consumption model in aircraft descent stage, a deep belief network combined with mutual information correlation analysis method is used to establish a fuel consumption model of aircraft descent stage based on flight parameters. The results show...
that compared with the traditional shallow neural network model, the accuracy of the model established by this method has been greatly improved and the error has been greatly reduced, which is of great significance.

References

[1] Dvid J Griggs, Joyce E Penner, David H Lister. Aviation and the global atmosphere[R]. Intergovernmental Panel on Climate Change, Cambridge University Press, 1999: 384.

[2] Bela P. Collins. Estimation of aircraft fuel consumption [J]. Journal of Aircraft, 1982, 11: 969–975.

[3] Wang Wei, Ning Dongfang, Zhang Jin. A genetic algorithm for the trajectory optimization of fuel-optimal flight based on Energy-state control [J]. Measurement & Control Technology, 2006, 25(1): 56-65 (in Chinese).

[4] Tian Husen, Xie Shousheng, Ren Litong, et al. The Fuel Consumption Model of Military Aircraft Based on Multiple Linear Regression [J]. 2014, 39(10): 1782-1785 (in Chinese).

[5] Baklacioglu T. Fuel Flow-Rate Modeling of Transport Aircraft for the Climb Flight Using Genetic Algorithms [J]. Aeronautical Journal -New Series, 2015, 119(1212): 173-183.

[6] Wu Wenjie, Hu Rong, Zhang Junfeng, et al. Research on aircraft energy-saving and emission-reduction of continuous descent approach based on BADA model [J]. Journal of Wuhan University of Technology (Transportation Science & Engineering), 2017, 41(4): 668-672 (in Chinese).

[7] Trani, A.A. Enhancements to SIMMOD: A Neural Network Post-processor to Estimate Aircraft Fuel Consumption [D]. Virginia: Polytechnic Institute and State University, 1997.

[8] Liu Jing. The aircraft fuel estimation model based on flight data analysis [D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2010 (in Chinese).

[9] Wang Chao, Zhou Xuanren, Wang Lei. Research on the estimation of air traffic fuel consumption based on trajectory data [J]. Journal of Air Force Engineering University (Natural Science Edition), 2018, 19(4): 25-30 (in Chinese).

[10] Tseng F M, Yu H C, Tzeng G H. Applied Hybrid Grey Model to Forecast Seasonal Time Series [J]. Technological Forecasting & Social Change, 2001, 67(2–3): 291-302.

[11] Tesmer M, Perez C A, Zurada J M. Normalized mutual information feature selection [J]. IEEE Transactions on Neural Networks, 2009, 20(2): 189.

[12] Shannon C E. A mathematical theory of communication [J]. ACM SIGMOBILE Mobile Computing and Communications Review, 2001, 5(1): 3-55.

[13] Hinton G E, Osindero S, Teh Y W. A Fast Learning Algorithm for Deep Belief Nets [J]. Neural Computation, 2006, 18(7): 1527-1554.