Research on Civil Aircraft Fuel Consumption in Cruise Phase Based on Least Square Support Vector Regression with Genetic Algorithm

Shuli Gong

College of Civil Aviation

Nanjing University of Aeronautics and Astronautics

June 22, 2018

Contents

Introduction

- Energy conservation and costs control are two particular issues for any airlines. Controlling fuel consumption is an effective way to solve both problems. A model is needed for the dispatch department of airline to forecast accurately fuel consumption during flight.
- The fuel consumption is related to many factors, such as flight data, aircraft performance, flight environment parameters, et al.
- The key to establish a reasonable fuel consumption model is acquiring aircraft performance data and flight environment parameters. Flight environment parameters include weather conditions, aircraft weight, flight status, flight data, etc.

Research on fuel consumption forecasting

- Schilling used neural networks to establish a model incorporating **flight performance data** ,such as lift, drag, thrust, and fuel flow rate to avoid manual operation and improve calculation speed. The flight environment parameters were not considered.
- Trani trained a neural network using a performance database combined with **flight data** to obtain fuel consumption. The aircraft performance or meteorological factors were not considered.
- Nikoleris et al. studied the fuel influencing factors in the taxi phase, analyzing fuel consumption from engine start to stop according to possible scenarios, such as acceleration, stop, and uniform motion. In this study, **aircraft performance parameters** were considered, but other factors, such as meteorological parameters and degradation of aircraft and engine performance were not considered.

object of this research

- The aircraft performance parameters,flight data and flight environment parameters were all considered to forecast fuel consumption.
- Due to the complexity of flight, the influence factors are not the same in different phase, only the cruise phase is considered here.

Aircraft Performance data

Lift coefficient

- The lift of aircraft is

$$L = C_L S_w \frac{1}{2} \rho V^2 \tag{1}$$

  where C_L is lift coefficient; S_w is wing area; V is true airspeed(TAS).
- When the aircraft is in level, the L equals to the gravity G.
  Larger lift coefficient, less TAS. So the thrust needed to maintain TAS is smaller. Given the specific fuel consumption(SFC),the fuel flow (FF) is reduced. **Fuel consumption is less** in the same period of time.

Aircraft Performance data

Drag coefficient

- The drag of aircraft is

$$D = C_D S_w \frac{1}{2} \rho V^2 \tag{2}$$

  C_D – drag coefficient;S_w – wing area;V is TAS
- **Smaller drag coefficient** ,less thrust. Given the specific fuel consumption(SFC),FF is reduced. **Fuel consumption is less** in the same period of time.

Aircraft Performance data

Lift to drag ratio (L/D)

Larger lift coefficient or Smaller drag coefficient, less fuel consumption.
So larger L/D, less fuel consumption.

Specific fuel consumption(SFC)

- SFC is defined as the quantity of fuel consumed by the engine within one hour to generate a specific thrust.
- Smaller SFC implies less fuel required, and hence fuel consumption rate is reduced.

Aircraft Performance data

- From above analyse, the aircraft performance data related to fuel consumption is lift coefficient, drag coefficient,Lift to drag ratio and SFC.In principle, dispatch departments should consider the parameters above to devise a flight plan.
- The performance data sheets available to dispatch departments do not directly correspond to above parameters and engine characteristics curves.
- For example, the B737-800’ s performance data sheet in cruise phase reflects the engine’ s FF for specific cruise altitude, total atmospheric temperature (TAT), fixed aircraft’s gross weight and airspeed. When the conditions change, the engine FF will also change.

performance parameter data sheet

Table 1: Cruise performance parameter range (CI=35)

Parameter	Minimum	Maximum
barometric altitude (ft)	1000	41000
gross weight (lb)	88105	167400
indication airspeed (knot)	230	337
total atmospheric temperature (° C)	-64.1	59.3
fuel flow (PPH)	1734	4147

Cruise performance parameter data sheet records aircraft performance data in terms of flight conditions, and is useful to forecast FF for each set of flight conditions. The fuel consumption can then be calculated by integrating FF over time.

Aircraft Flight data & Environment data

QAR (Quick Access Recorder) records environment data and flight data corresponding to aircraft flight performance,such as airspeed, barometric altitude, TAT, Mach, heading, AOA, aircraft gross weight, angle of pitch, angle of roll, vertical acceleration, longitudinal acceleration, wind speed ,wind direction ,etc.

SVR & LSSVR

- Support vector machine (SVM) is a machine learning method based on statistical learning theory. It can solve practical problems such as small sample size, nonlinear, and high dimensional pattern recognition.
- Vapnik and others introduced the insensitive loss function,ε, which extended SVM to regression estimation for non-linear systems (support vector regression (SVR)). Due to its excellent learning ability, SVR has been widely used in many fields, such as non-linear regression estimation, non-linear system modeling, etc.
- SVR has slow convergence. Least square support vector regression (LSSVR) inherits the advantages of SVR, simplifies the modeling process, and increases convergence speed.

Theory of LSSVR

- Compare with standard SVR, LSSVR tries to minimize primal cost function subject to equality constraints instead of inequality ones. Therefore, LSSVR solves a set of linear equations instead of computational cost quadratic programming problem.
- The LSSVR optimization problem is

$$minJ(w,e) = \frac{1}{2}w^T w + \frac{1}{2}C\sum_{i=1}^{n} e_i^2 \tag{3}$$

$$s.t. \quad y_i = w^T \phi(x_i) + b + e_i, i = 1, 2, \cdots, n \tag{4}$$

- where n is the total number of data points; $x_i \in R^n$ is the input vector, and $y_i \in R$ is the corresponding output vector;J is the loss function;$\phi(x)$ is the non-linear transformation function; w is the weight vector; b is the bias term; $e_i \in R$ is the slack variable; and $C > 0$ is the regularization parameter.

Transforming (3) into the dual form using Lagrange multipliers $a_i \in R$

$$L(w, b, e, a) = \frac{1}{2} w^T w + \frac{1}{2} C \sum_{i=1}^{n} e_i^2 - \sum_{i=1}^{n} a_i \{ w^T \phi(x_i) + b + e_i - y_i \} \tag{5}$$

Applying the Karush – Kuhn – Tucker’ s (KKT) conditions to obtain the following equation, w and e can be eliminated, and get

$$\begin{bmatrix} 0 & I^T \\ I & \Omega + C^{-1} I \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{6}$$

where

$$\Omega_{kl} = \phi(x_k)^T \phi(x_l) = K(x_k, x_l), \qquad k, l = 1, 2, \ldots, n \tag{7}$$

K is called the kernel function .

LSSVR estimation function is

$$y(x) = \sum_{k=1}^{n} a_k K(x, x_k) + b \tag{8}$$

Radial basis function (RBF)is selected as the kernel function.

$$K(x_i, x_j) = exp(-\frac{\|x_i - x_j\|^2}{2\sigma^2}) \tag{9}$$

C and σ are user-determined parameters, the selection of them plays an important role in the performance of LSSVR.

Optimization of LSSVR parameters with GA

Genetic Algorithm(GA) is a random search algorithm that simulates natural selection and genetic process with implicit parallelism and strong global search capability. Therefore, it can search for the global optimal efficiently, and would be an effective way to optimize LSSVR parameters.

The optimization process is shown in Fig. 1.

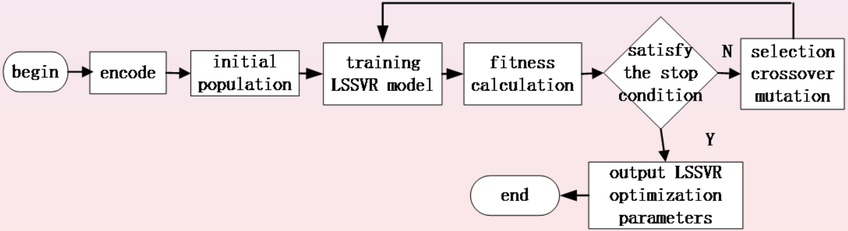


Figure 1: chart of optimization process

Process of optimizing LSSVR parameters with GA(1)

- Determine the range of possible values for C and σ for practical applications;
- Encode the parameters in binary system and generate the initial population randomly;
- Using the the mean square relative error to calculate fitness, i.e.

$$Fitness = \sqrt{\frac{1}{l}\sum_{i=1}^{l}(\frac{(y_i - \hat{y}_i)}{y_i})^2} * 100\% \tag{10}$$

where l is the number of samples, y_i and $\hat{y}_i$ are the actual and validation values respectively.

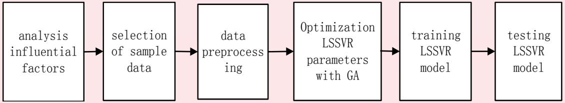
Process of optimizing LSSVR parameters with GA(2)

- Implement genetic operations to generate offspring. The genetic operation includes selection, crossover, and mutation operations.
- Judge if the process has reached a stop condition. The stop condition is a maximal generation. If the stop condition is not satisfied, the process returns to step 3. Otherwise, individuals with minimal fitness value are taken as the optimal solution, and the evolutionary process stops.

Process of forecasting fuel consumption with LSSVR-GA

The flow chart to establish forecasting civil aircraft fuel consumption model during cruise phase based on LSSVR-GA is shown in Figure 2.

- Determine the influential factors , selection of sample data and preprocessing the data;
- Optimize model parameters with GA;
- Train and test the LSSVR model.


analysis influential factors
selection of sample data
data preprocessing
Optimization LSSVR parameters with GA
training LSSVR model
testing LSSVR model

data preparation

- The proposed LSSVR-GA model is trained using barometric altitude, gross weight, airspeed, and TAT, as recorded by QAR. The model output is engine FF, corresponding to the initial performance data sheet.
- The model is tested on further QAR recorded data, where the output was the engine FF corresponding to the initial performance data sheet integrated over time.
- The sample data is obtained from a commercial B737-800 flight data on the route from Zhengzhou to Yinchuan. There are 30 groups of flight data in total, 25 are selected as training data, and the other 5 as test data.

data processing

There are five parameters related to the proposed model input and output, with significantly different dimensionalities. The original of every parameter should be normalized over $[0, 1]$ using this formula:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{11}$$

Training and Testing the LSSVR-GA Model

initial conditions

initial population size = 10, maximal generation = 100, crossover probability =0.8, mutation probability =0.05, $C = (0, 200)$ and $\sigma = (0, 100)$.

The GA toolbox GAOT and MATLAB toolbox LS-SVM software are used.

LSSVR parameters

With increasing iteration, fitness improves rapidly at first, and then tends to stabilize. When $\mathbf{C} = \mathbf{176.38}, \sigma = \mathbf{1.39}$, and iteration is 60, fitness converged to 1.154.

Simulation Results

- In order to verify the forecasting performance of LSSVR-GA model, a standard SVR model and a three layer BP neural network (BPNN) are also used to forecast the fuel consumption with the same train and test data.
- From the prediction accuracy and computational speed to compare the three model. The prediction FF is shown in Fig.3. Average RMSE and computational speed are shown in Table 2.
- In Fig.3,there are only the result curves of BNPP and LSSVR-GA,except the standard SVR.Because the results of SVR and LSSVR-GA are indistinguishable.
- In table 2,the values of 5 flights' actual fuel consumption,three models' forecasting fuel consumption, average RMSE and computational time are shown.

Simulation results(1)

- Black curve is the actual fuel consumption recorded by QAR.Green and red curves are the predict fuel consumption with BNPP and LSSVR-GA respectively.
- At most of the cruise time, the accuracy is controlled in a steady level. With regard to FF's sudden change , two models have large error.

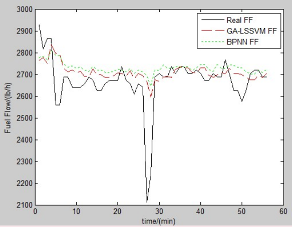


Figure 3: fuel consumption forecasting results

Simulation results(2)

		SVR model			LSSVR-GA model			BNPP model		
Flight NO.	Actual Value(kg)	Value(kg)	RMSE(%)	Time(s)	Value(kg)	RMSE(%)	Time(s)	Value(kg)	RMSE(%)	Time(s)
1	2658	2475	6.8	13.12	2475	6.8	5.73	2473	6.9	6.09
2	2527	2416	4.4	13.17	2428	3.9	5.81	2414	4.8	6.37
3	2830	2654	6.2	12.75	2647	6.4	5.12	2607	7.9	5.42
4	2976	2703	9.2	12.97	2719	8.6	5.47	2704	9.1	6.23
5	2453	2396	2.3	11.78	2403	1.6	4.98	2365	3.5	5.13

- All three methods have high accuracy in forecasting fuel consumption.
- BPNN tends to become trapped into local optima and over learning, which can produce individual large errors.
- SVR has high forecasting accuracy, but computational speed is low.
- Compared to BNPP and SVR,LSSVR-GA is superior because it provides excellent accuracy and desirable speed.

reason of error

- Limitation of forecasting model;
- The output of model is corresponding to the initial aircraft performance data sheet, without considering the degradation of aircraft performance;
- The number of sample data is too small,especially for BNPP model.

Conclusion

- Proposed a fuel consumption forecasting model for cruise phase based on LSSVR-GA, where GA was employed to optimize LSSVR parameter.
- Four factors were identified to affect fuel consumption in cruise phase: barometric altitude, gross weight, airspeed, and total atmospheric temperature, and these formed the model inputs. The model output was engine fuel flow, corresponding to the aircraft initial performance data sheet.
- Flight data from a B737-800 on the Zhengzhou to Yinchuan route was chosen as the sample data to train and test the model. Experimental results show that the proposed LSSVR-GA model provides equivalent accuracy to traditional SVR and BPNN models, and agrees well with actual data, but avoids being trapped in local optima (as per BPNN), and offers significantly faster computation than SVR.

Further research

- The accuracy of the proposed method needs to be improved.
- With increasing flight time, aircraft aging, engine performance attenuation, and flight environment changes will cause increased fuel consumption. Therefore, the focus of the future work is to dynamically regulate the model to improve forecasting accuracy for real and in-service aircraft.

Thank you!