DOI: 10.13009/EUCASS2017-100

A comparative study of mathematical modeling methods for rocket aerodynamic data

*Qian Wei-qi***, Yang Jun**, and Zhou Yu****
**State Key Laboratory of Aerodynamics*
Mianyang Sichuan 621000, China
***China Aerodynamics Research and Development Center*
Mianyang Sichuan 621000, China

Abstract
Mathematical modeling with aerodynamic data plays an important role in the evaluation of flight performance simulation. In this paper, three kinds of modelling methods, the Orthogonal Least Squares method, the Kriging model, and the classification and regression tree method, are utilized for the aerodynamic data modelling of a rocket, which belongs to the axial symmetric winged aircraft. Also a training sample selection method based on the idea of uniform design with Genetic Algorithms is developed. After applying these methods for a test case, it can be seen that, there is a significant improvement of the prediction ability of the model thanks to the adoption of the new GA-based training samples selection method. All these three modelling methods can do a good job for the mathematical modelling of test case, the modelling results of OLS is of high accuracy and good generalization capability, but it's only suitable for a certain kind physical problem. On the other hand, both the Kriging model and the CART method have a good universality, but their modelling accuracy and generalization capability are greatly influenced by the training samples.
Keywords: Aerodynamic data modelling, Orthogonal Least Squares method, Kriging model, Classification and regression tree method, Selection of training samples

1. Introduction

Aerodynamic modelling is the basis of the performance evaluation for the aircrafts. As far as the winged axial symmetric aircraft of rocket is concerned, its aerodynamic characteristics is a six-dimensional function of Mach number, attack angle, body rolling angle and three control surface deflection angles. So it's hard to get all the aerodynamic data in the full flight envelop by CFD or wind tunnel experiments. When the attack angle is small, this problem can be abated by establishing a linear mathematic model using aerodynamic derivatives to get the aerodynamic characteristics in different Mach number, different attack angle and different control surface deflection angles. But at present, with the improvement of the maneuverability of modern rocket, the attack angle and the control surface angle becomes large and the nonlinear aerodynamic characteristics comes out, which means that the linear method of aerodynamic modelling can not satisfy the engineering requirement and it's necessary to develop some new method to build the multivariate nonlinear aerodynamic model from wind tunnel experiment or CFD data[1-3].

There are mainly two kinds of modelling methods for nonlinear aerodynamic characteristics. First is the rational modelling method based on physical mechanism, such as the polynomial function model, and the differential/integral equation method which is exemplified by the unsteady aerodynamic under the high-angle of attack. Second is Response Surface Model(RSM) based on training samples learning, such as the neutral network model, the fuzzy logic model, and the Kriging model. By comparison, the RSM method fits the sample data better than the physical mechanism modelling method, but the terms in the physical mechanism model are physically meaningful and having better generalisation capability. In this paper, three modelling methods, i.e., the Orthogonal Least Squares(OLS) method, the Kriging method, and the Classification and Regression Tree(CART) method, are utilized for the construction of model with the aerodynamic data of a rocket and both the advantage and disadvantage of these models are analyzed.

DOI: 10.13009/EUCASS2017-100

2. Orthogonal least squares method

For rocket, the aerodynamic model often can be expressed as a trigonometric series function of the roll angle and the control surface deflection angles because there exits some geometric symmetry nature in the data. Ref.[7] build such model that for certain Mach number and attack angle, the aerodynamic coefficient could be expressed as the sum of two parts, one is the aerodynamic coefficient with control surfaces at the neutral position, the other is the influence of control surface deflection. Taken the longitudinal aerodynamic coefficient as an example, it could be written as,

$$F = a_0 + \sum_{i=1}^{m_1} a_i \cos(n_i\phi) + \sum_{j=1}^{m_2} a_{m_1+j}[f_j(\delta_x,\delta_y,\delta_z)\cos(n_j\phi)] + \sum_{k=1}^{m_3} a_{m_1+m_2+k}[g_k(\delta_x,\delta_y,\delta_z)\sin(n_k\phi)] \quad \textbf{(1)}$$

where ϕ is roll angle; a_i is coefficients of series; n_i、n_j、n_k are positive integers; m_1 is the number of trigonometric function terms without control surface influence; m_2 and m_3 are the number of cosine and sine base function terms taking control surface deflection influence into accounts; f_j, g_k are the polynomial function of control surface deflections. And more details can be found in Ref.[7].

The basic idea of the OLS method is that for the functional form of Eq.(1), the terms with little influence on the modelling result are abandoned and the least square method is adopted to find the coefficients of the series which are left. The general procedure of the OLS is as follows. In the Euclidian space spanned by a group of vectors $\vec{x}_1$, $\vec{x}_2$,..., $\vec{x}_m$, the parameter estimation problem can be written as,

$$\vec{y} = \theta_1\vec{x}_1 + \theta_2\vec{x}_2 + \cdots + \theta_m\vec{x}_m + \varepsilon \quad \textbf{(2)}$$

It can be turned into the equivalent form of

$$\vec{y} = g_1\vec{w}_1 + g_2\vec{w}_2 + \cdots + g_m\vec{w}_m + \varepsilon \quad \textbf{(3)}$$

Where $\vec{w}_1$,..., $\vec{w}_m$ are the orthogonal basis of $\vec{x}_1, \vec{x}_2$,..., $\vec{x}_m$, and they can be obtained from $\vec{x}_1, \vec{x}_2$,..., $\vec{x}_m$ by Gram-Schmidt orthogonalization procedure. From Eq.(3), the contribution of each term to the output can be justified and filtered by the following criterion.

$$e_i = \frac{g_i^2\langle\vec{w}_i,\vec{w}_i\rangle}{\langle\vec{y},\vec{y}\rangle} \quad \textbf{(4)}$$

Where "<·>" denotes the inner product and if e_i is less than the threshold of σ, the i_th term can be ignored in the regression.

3. Kriging method

This model has its origin in mining and geostatistical applications involving spatially and temporally correlated data and combines a global model plus localized departures. In fact, the Kriging model is a special case of the Gaussian Process(GP) model. The kernel idea of GP is treating the training samples as a Gaussian process and every two samples satisfies the joint Gaussian distribution, and when a new sample is presented, this probability function can be used to estimate the output. Especially, the model is specified by its mean function, and a covariance function (a function which looks at the covariance between responses at a pair of sample data points). The parameters that control the covariance function are called hyperparameters which can be decided by Bayesian inference method.

In Kriging model, when the number of the training samples is n_s, and the input variable of the sample, $\vec{x}$, is n-dimensional, the output variable of all the samples $\vec{y}$ can be approximately modeled as,

$$\hat{\vec{y}} = \beta + \vec{r}^T(\vec{x})\boldsymbol{R}^{-1}(\vec{y} - \vec{f}\beta) \quad \textbf{(5)}$$

where $\hat{\vec{y}}$ is a n_s-dimensional vector, and $\vec{f}$ is also a n_s-dimensional vector whose element are all equal to 1. And

$$\vec{r}^T(x) = \left[R(\vec{x},\vec{x}^1), R(\vec{x},\vec{x}^2),\dots, R(\vec{x},\vec{x}^{n_s})\right]^T ; \quad R\left(\vec{x}^i,\vec{x}^j\right) = \exp\left[-\sum_{k=1}^{n}\theta_k\left|x_k^i - x_k^j\right|^2\right] \quad \textbf{(6)}$$

where the superscript and subscript of $\vec{x}$ denotes the sample index and the vector component, $R(\vec{x}^i,\vec{x}^j)$ is the covariance function and θ_k are hyperparameters. Form Eq.(5), the estimated value of $\hat{\beta}$ can be obtained as,

$$\hat{\beta} = \left(\vec{f}^T\boldsymbol{R}^{-1}\vec{f}\right)^{-1}\vec{f}^T\boldsymbol{R}^{-1}\vec{y} \quad \textbf{(7)}$$

And the estimated variance value of $\hat{\sigma}^2$ is,

$$\hat{\sigma}^2 = \left[(\vec{y} - \vec{f}\hat{\beta})^T\boldsymbol{R}^{-1}(\vec{y} - \vec{f}\hat{\beta})\right]/n_s \quad \textbf{(8)}$$

DOI: 10.13009/EUCASS2017-100

It is the hyperparameters, θ_k, that are optimized to train the model. This is performed by maximizing the following logarithmic likelihood.

$$\max_{\theta_k>0} \Phi(\theta_k) = -\frac{1}{2}\left[n_s \ln(\hat{\sigma}^2) + \ln|\mathbf{R}|\right] \tag{9}$$

This is a nonlinear optimization problem, and when the values of θ_k are optimized, the Kriging model is then constructed. The output of a new input $\vec{x}_{pre}$ is

$$y_{pre} = \hat{\beta} + \vec{r}^T(\vec{x}_{pre})\boldsymbol{R}^{-1}(\vec{y} - \vec{f}\hat{\beta}) \tag{10}$$

where the vector of $\vec{r}^T(\vec{x}_{pre})$ is calculated with Eq.(6).

4. Training samples selection method based on GA

In the aforementioned two modelling methods, training samples are needed to build the model and decide the parameters and the selection of training samples out from a significant number of candidate samples plays an important role on the modelling result. It is obvious that it would be better that the training samples be uniformly distributed in the parametric space, so in this paper a training samples selection method based on uniform design is put forward. In this method, different combinations of training samples are generated from the candidate samples and the combination having the maximum index of uniformity can be selected. This is a combinatorial optimization problem and can be solved with Genetic Algorithms(GA). When the total number of the candidate samples is M, and the number of selected training samples to build the model is N, the detail of the algorithm is given as follows.

First, using the binary coding rule, define one possible combination of training samples as a “chromosome” whose length is equal to N and composed of values of $g_i \in [0,1]$, (i=1,N). Every g_i corresponds to the integer index m_i(i=1,N) of the selected training samples in the total candidate set. The relationship between g_i and m_i is

$$m_i = 1+\mathrm{Int}((M-1)*g_i) \tag{11}$$

Note that either two selected training samples in the chromosome should be different,

$$m_i \neq m_j \text{ when } i \neq j. \tag{12}$$

Furthermore, in order to adopt the GA, for every chromosome, the fitness can be defined as the uniformity of the N samples in the chromosome, and the following Symmetric Deviation in Ref.[9] is often taken as the measurement of the uniformity

$$SD_2 = \left[\left(\frac{4}{3}\right)^s - \frac{2}{N}\sum_{k=1}^{N}\prod_{i=1}^{s}\left(1+2x_{m_k,i}-2x_{m_k,i}^2\right) + \frac{2^s}{N^2}\sum_{k,l=1}^{N}\prod_{i=1}^{s}\left(1-\left|x_{m_k,i}-x_{m_l,i}\right|\right)\right]^{\frac{1}{2}} \tag{13}$$

Where s is the dimension of the sample and $x_{m_k,i}$ is the i_th component of the sample indexed m_k, and its value has been scaled to the range of [0,1].

Second, generate the initial population consisting of P individuals. Every individual corresponds to a chromosome.

Third, calculate the fitness of every individual in the population with Eq.(12) and perform the duplication and selection operation on this population by weighted roulette wheel method, and a new population of P individuals can be obtained.

Fourth, perform the one-point crossover and mutation operations on the population in the third step and generate new individuals and population. Note that in order to guarantee the constraint of Eq.(13), the new individuals after crossover and mutation should be checked and if the constraint is not satisfied, the crossover and mutation operations are repeated.

Last, find the individual with the largest fitness value, and decide the population evolution is convergent or not. If convergent, stop the evolution, otherwise, return to the third step to continue the evolution.

5. Classification and Regression Tree(CART) method

CART is an interesting and effective non-parametric classification and regression algorithm which has been more and more widely used in the aeronautical research field[10, 11]. Different from traditional statistic method, CART constructs the binary regression tree from training samples and uses the tree to carry out the prediction. The main structure of the regression tree is shown in Fig.1. Starting with the root node containing all the data points, CART carries out recursive binary splitting of the data. The split criteria are often of the form Xi<Ti where Xi is a particular

DOI: 10.13009/EUCASS2017-100

predictor variable and Ti is the split point, or the threshold value. Points in a node satisfying the split criterion go into the left child node and the others go into the right child node. This split at each node is a locally optimal split, chosen so as to maximally reduce the weighted sum of the mean squared errors of the resulting nodes. The sum is weighted by the fraction of the observations going in the left child and the right child nodes. The nodes which do not further split are known as leaf nodes and every leaf node corresponds to a training sample data. In practice, the construction of the regression tree contains two steps, first step is using part of the training samples to build the tree, and the second step is trim, adding or merging the nodes to optimize the regression performance of the tree with other training samples.

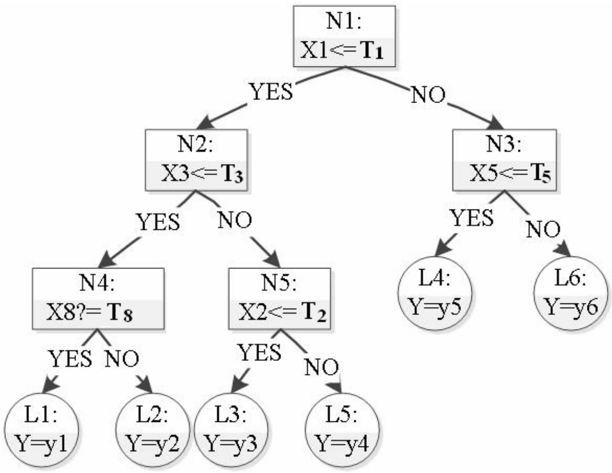


Figure 1: Example of regression tree of CART method

6. Test case

A typical rocket shown in Fig.2 is investigated in the test case. The flight condition is chosen as Mach number M=1.5, and attack angle α=6°, and four control surfaces deflect schematically to generate 216 combinations of three equivalent rolling, yawing, and pitching control surface deflections of δx= 0°, 5°, 10°; δy= -15°, -10°, -5°, 0° , 5°, 10°, 15°, 20°; and δz= -15°, -10°, -5°, 0° , 5°, 10°, 15°, 20°. Since every combination corresponds to 8 body roll angles, ϕ= -135°, -45°, 0°, 22.5°, 45°, 67.5°, 90°, 180°, the normal force coefficients of totally 216×8=1728 states are calculated with DATCOM software[8] and the calculated results are used as the experimental data for the following aerodynamic modelling.

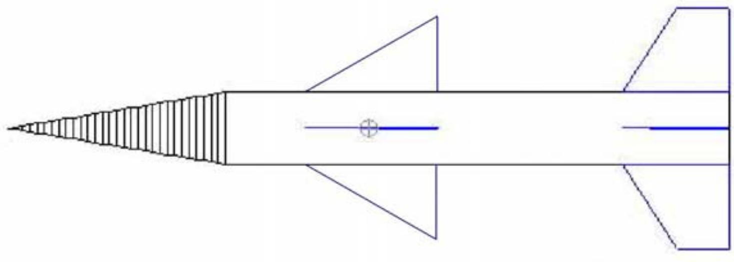


Figure 2: A typical rocket geometry

Two methods are used to generate the training samples. The first method is conventional and regularly selecting the states involving "control surfaces at the neutral position" state, "identical control surface deflections in rolling, yawing, and pitching channel" state, and "different deflection angles in rolling, yawing, and pitching channel" states such as "δ_x= 0°, δ_y= 0°, δ_z= 0°, ϕ=0°, 22.5°, 45°, 67.5°, 90°", "δ_x= 0°, δ_y=δ_z=5°,10°,15°, ϕ= -135°,-45°,45°", "δ_x=5°, δ_y=5°, δ_z=10°, ϕ=0°", and so on. The second method is the GA-based selection method presented in Section 4. For OLS and Kriging method, 50 samples are selected as training samples from 1728 candidates, and the left 1678 samples would be used for prediction. The prediction error is defined as,

$$E = \sqrt{\sum_{i=1}^{1678}\left[C_{A\mathrm{model}}(\delta_{xi},\delta_{yi},\delta_{zi},\phi_i) - C_{A\mathrm{exp}}(\delta_{xi},\delta_{yi},\delta_{zi},\phi_i)\right]^2 \Big/ 1678} \tag{14}$$

where the subscripts "model" and "exp" denote the model prediction value and the experimental value respectively.

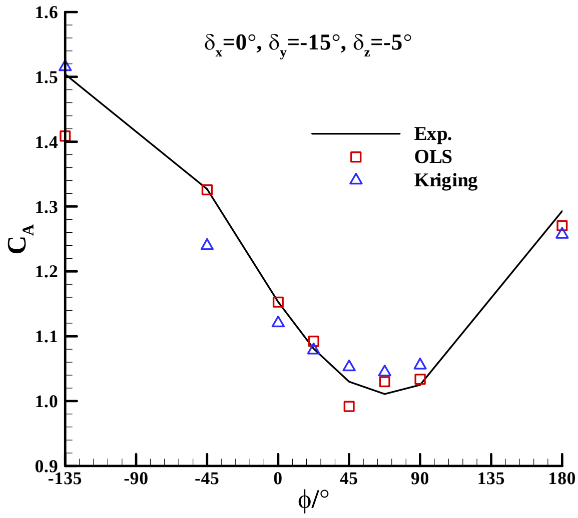
DOI: 10.13009/EUCASS2017-100

For OLS method, a trigonometric series function involving 50 terms are used as the modelling basis of OLS, and the model term contribution threshold of σ is set to be σ=0.001. For Kriging model, GA is also used to carry out the optimization of hyperparameters in Eq.(9).

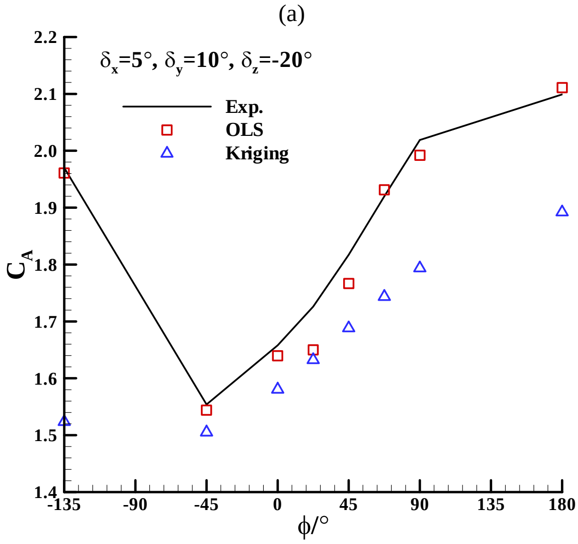
Table 1 The comparison of prediction error for OLS and Kriging model

Modelling method	*E* (Regular training samples selection method)	*E* (GA-based selection method)
OLS	0.31025	0.04602
Kriging model	0.37347	0.19241

The comparison of prediction error for OLS and Kriging model corresponding to different training sample selection methods are given in Table 1. And for the GA-based training samples selection case, the model prediction results by OLS and Kriging model for two sets of control surfaces deflection are compared in Fig.3, in which “OLS”, “Kriging”, “Exp.” denotes the model prediction of OLS method, the model prediction of Kriging model, and the experimental value.



(a)



(b)

Figure 3: Comparison of model prediction results for two sets of control surfaces deflection

From the table and the figure, it can be seen that, first, when the GA-based training samples selection method is adopted, there is a significant improvement of the prediction error. Second, although the Kriging model can fit all the training samples, its generalization capability is greatly influenced by the complexity of the physical problem, i.e., under some circumstances the Kriging model’s prediction fits the experimental value well, such as Fig.3(a), while

DOI: 10.13009/EUCASS2017-100

under other circumstances the prediction is not so good, such as Fig.3(b). Theoretically, the more training data used, the better the Kriging model will be, but it can be seen from Eq.(5) that if training data number increase, the computation of covariance matrix inversion would be much more complicated and computationally expensive. On the other hand, for the OLS, because the basis functions of the model reflect certain physical mechanism, OLS method has a good generalization capability and its model prediction results are generally better than the Kriging model.

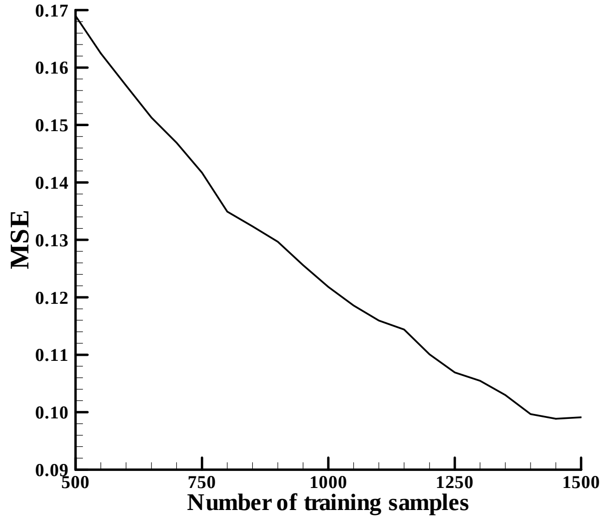


Figure 4: Influence of number of training samples on modelling result

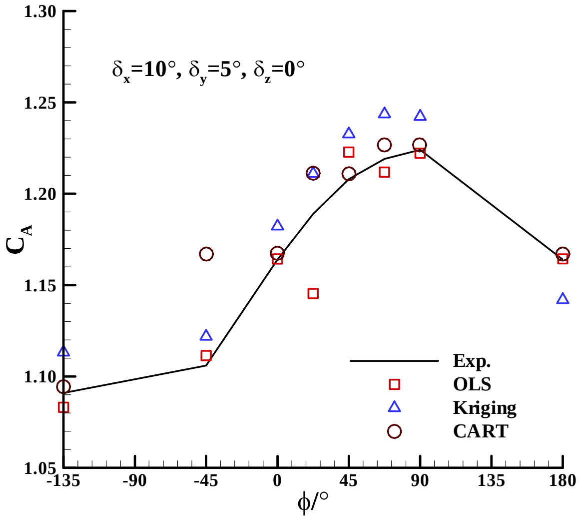


Figure 5: Comparison of model prediction results

For CART method, at first, the influence of the number of training samples on the modelling results are illustrated in Fig.4, the horizontal axis denotes the number of training samples to build the regression tree in CART for this test case, and the vertical axis denotes the Mean Square Error(MSE) between the CART model prediction and the experimental value for all the 1728 samples. It can be seen that the more training data used in CART, the lower is the MSE. A set of CART model prediction result with the training samples numbers of 250 is shown in Fig.5 and compared to the aforementioned model prediction result of OLS and Kriging model. The results show that the result of CART method is agreeable with OLS and Kriging model prediction and the experimental value, and is feasible for the aerodynamic data modelling.

7. Conclusion

In this paper, three modelling methods, OLS method, the Kriging method, and the CART method, are briefly

DOI: 10.13009/EUCASS2017-100

introduced and utilized for the aerodynamic modelling of a rocket. From the test case, the following conclusions can be drawn.

Firstly, all these three methods can do a good job on the mathematical modelling with the axial force data of the rocket. a prior information about the model structure should be known in advance when using the OLS method, while such information is not needed for the modelling of GP or CART, that's to say, the latter two methods can be widely applied to more cases.

Secondly, training sample selection may play an important role on the modelling result of OLS and Kriging model, when the training samples are uniformly distributed in the parametric space, both methods can get certain good results when the training sample number is not so large. On the other side, the performance of the CART method relies much on the number of training samples, and when the number of samples increase, there is a significant improvement of the CART model's prediction ability.

Thirdly, as for the modelling accuracy, the OLS is better than the Kriging model because the generalization capability of OLS is better than the Kriging model. If more training sample data is used in the Kriging model, its generalization capability may increase but lead to heavier computations of large covariance matrix inversion. As mentioned before, the accuracy and generalization capability of CART model improves with the increase of training samples, which is especially effective in "Big data" case.

Lastly, lots of matrix operations are processed when using the OLS and Kriging model. Especially when dealing with larger amount of training sample data, these two methods both play poor performance due to the possible singularity of the matrix. But the CART method can works well because there is no need of matrix operations in it, which means that it is more convenient and efficient to use the CART method in engineering practices.

Acknowledgements

The authors are grateful for the support provided by the National Natural Science Foundation of China (Grant No.11372338).

References

[1] Grauer J. A., and E. A. Morelli. 2014. A generic nonlinear aerodynamic model for aircraft. AIAA-2014-0542. AIAA

[2] He, K. F., Wang, W. Z., and W. Q. Qian. 2004. Mathematic modeling for the missile aerodynamics with tail-wing according to wind-tunnel test results. *Experiment and Measurements in Fluid Mechanics*. 18(4):62-66. (In Chinese)

[3] He K. F., Qian W. Q., Zhang Y., and W Z Wang. 2014. Application of partial least squares regression method in modeling of aerodynamic data. *Journal of Astronautics*. 35(3):277-282. (In Chinese)

[4] Fu, J. M. 2005. Three dimensional aerodynamic mathematical model for tactical missiles with jet steering. *Aerospace Shanghai*. 4:13-18. (In Chinese)

[5] Zheng K. P., and H. F. Li. 2013. Missile aerodynamic coefficient model selection method using OLS algorithm. *Command Control＆Simulation*. 35(4):75-76(+80). (In Chinese)

[6] Liu X. L., Zhang K. L., and B. Guo. 2010. Research on Gaussian process surrogatemodel and its application to rocket aerodynamic analysis. *Journal of Solid Rocket Techno logy*. 33(5):486-490. (In Chinese)

[7] Simpson T W, et al. 2001. Kriging models for global approximation in simulation- based multidisciplinary design optimization. *AIAA Journal*. 39(12):2233-2241.

[8] Zhou M., and S. D. Sun. 1999. Genetic algorithms and its applications. National Defense Industrial Press, Beijing. (In Chinese)

[9] Fang K. T. and C. X. Ma. 2001. Orthogonal and uniform design of experiments. Science Press, Beijing. (In Chinese).

[10] Chati Y. S.and H. Balakrishnan. 2016. Statistic modeling of aircraft engine fuel flow rate. ICAS-2016-0619. ICAS.

[11] Yang J. 2016. Preliminary study of methods using in modeling and data mining for aerodynamic data. Master Thesis, China Aerodynamics Research and Development Center, Mianyang, China. (In Chinese)

[12] Auman, L. 2008. MISSILE DATCOM User's Manual-2008 Revision. AFRL-RB-WP-TR-2009-3015.