32nd Congress of the International Council of the Aeronautical Sciences
ICAS 2020 SHANGHAI
September 6-10, 2021
Pudong Shangri-La, Shanghai, China

A manufacturing cost estimation model with reliability for civil aero-engine

Xiaocui-Li[1], Huifeng-Wei[2], Jingmin-Wu[3], Fang-Liu[4]

[1,2,3,4]Aviation Industry Development Research Center of China

Abstract

In this article, we built a production cost estimation model based on the limited information of engine in development stage. In the modeling process, the reliability factor is taken into account. Therefore, the model can not only realize the production cost estimation in the development stage, but also reflect the quantitative relationship between reliability and production cost, which provides a more comprehensive support for the multi-objective trade-off decision making in the development stage of civil aero engine.

Keywords: estimation model manufacturing cost reliability civil aero-engine

1. Introduction

At present, civil aero-engine not only needs to fulfill the technology and function requirements from customer, but also to keep a competitive price. The development and manufacture process of aero-engine has its own characteristics, so we need to research the theories and methods for balancing the technology, cost and schedule of aero-engine particularly, which has become one of the vital research topics in the civil aircraft area now.

Recent years, the increasing demand for civil aero-engine reliability is bringing a greater pressure on cost reducing work. It's receiving more and more engineers' attention to achieve a higher reliability as well as a lower cost, which is also a difficult problem for civil aero-engine integrated balancing work. As a critical part of civil aero-engine LCC (life cycle cost), manufacturing cost is the key for civil aero-engine price decreasing and influences the civil aircraft business success in a degree. As we all know, the manufacturing cost is mainly decided at development phase, when the conceptual project is designed. At that phase, the information we can get are limited, making the balancing work between manufacturing cost and reliability difficult.

Under the research requirements and limitations above, we devote to build a manufacturing cost estimation model based on limited technology and function information at the development phase while considering the reliability factor. This model makes it possible for estimating the manufacturing cost as well as reflecting and analyzing the relationship between reliability and manufacturing cost. Therefore, this model can help balancing between the reliability and manufacturing cost while support conceptual project improvement and selection for decision maker in a degree. In a long course, it can be an effective tool for civil aero-engine integrated balancing work.

2. Research Actuality

Since the 1980s, the United States has formulated a series of aero-engine development plans, dedicated to improving the overall performance of aero-engines while achieving the economic affordability. Including the Integrated High-Performance Turbine Engine Technology (IHPTET) plan, the general and affordable advanced turbine engine (VAATE) plan [1], etc. IHPTET plans to invest 5 billion US dollars to achieve a "push-to-weight ratio (power-to-weight ratio) increase of 100%-120%, a 30%-40% reduction in fuel consumption, a 35%-60% reduction in production and maintenance costs, and affordability in three stages. The target is to increase the affordability on the basis of the F119 engine. After the great success of the IHPTET program, the United States launched the VAATE program. It is the world's first aero-engine technology development program with the goal of affordability. The focus of development has shifted from improving performance and reducing costs to improving affordability. In addition, the European Advanced Core Military Engine (ACME) program, led by the United Kingdom, and the "Affordable Near-Term Low Pollution (ANTLE)" program led by Rolls-Royce have put forward specific goals to improve performance and reduce costs. A variety of effective measures have been taken in aerodynamics, thermodynamics, material technology, structural strength and control.

The study of cost estimation models is an important part of the research on aero-engine affordability. In the process of effectively advancing and implementing aero-engine development plans abroad, a variety of aero-engine cost estimation models have been established to lay a technical foundation for the effective implementation of the development plan. Aiming at the production cost estimation problem in the development stage, the American RAND Corporation established the DAPCA-IV engine production cost estimation model with the maximum thrust, the maximum Mach number and the turbine inlet temperature of the engine as parameters. The U.S. Air Force has established a molar factor estimation model using aero engine turbine inlet temperature, maximum fuel consumption, and air flow as parameters. The Central Aviation Engine Research Institute of Russia established an engine production cost estimation model with parameters such as engine afterburner thrust, maximum take-off thrust, year of supply, and engine generation. P&W has established a method for estimating the manufacturing cost of engine components. It can be seen from the above that in the development stage, due to the lack of data and information about the engine, the parameter method is generally used in each research structure, and the technical indicators of the effective engine are selected as the parameter factor to establish the production cost estimation model.

Domestic research work on cost analysis began in the 1980s. Relevant documents were released, including GJB1364 "Equipment Cost-Effectiveness Analysis" and "Comprehensive Demonstration Report Compilation Guide." Proposed related concepts, including the introduction of earned value management (EVM), the work breakdown structure (WBS), and establishment of the "four characteristics in one" (Applicability, comfort, affordability, commonality and serialization), focusing on direct operating costs (DOC) and lifetime costs (LCC)". However, our economic work for aero engine turbine started late, with less data and experience accumulation. In terms of economics and cost analysis, a lot of research work has been done in this regard, while most of the work is qualitative analysis, and the quantitative analysis is insufficient.

Based on the research results at home and abroad, we found that, for the production cost estimation in the development stage, the domestic cost estimation model developed late and there are fewer models, and mainly draws on the experience of foreign modeling. The cost estimation models established abroad generally adopt the parameter method, and all use technical indicators as parameters. However, in the current domestic civil aviation engine, it is required not only to consider the technical indicators and production costs, but also to consider the reliability of the engine in the development stage. This requires us to study the relationship between production costs, technical indicators and reliability, so as to achieve multi-objective trade-offs in the development phase.

3. Modeling ideas

The model building process including cost factor choice, data collection, modeling, model application and so on. We collect the information of typical civil aero-engine first, and then do cost factor diming and correlation analysis on the date collected. We get a critical cost factor group-including reliability-through the work above, and finally succeed an estimation model with scientific modeling methods on the cost factor group. The cost factors in relation with the manufacturing are various and complicated, so it hard to analysis. Meanwhile, the civil aero-engine information is difficult to obtain.

All of those above make the realization of this research a significant achievement in the civil aircraft field at home.
Based on the experience on building the production cost estimation models at home and abroad, and combining the characteristics on civil aero engine development phases, we still chose to use the parameter method concept to establish a civil aero engine production cost estimation model on the development phase. As we all know, the key to parameter method is the selection of parameters and the determination of model form. In this part, we explain in detail the idea of determining the parameters of the cost estimation model and the basic form of the model.

3.1 parameter selection

There are many factors that affect the cost of engine production, but in the engine development stage, the information of the engine that can be obtained is not abundant. In the process of constructing the production cost estimation model, how to choose the parameters scientifically and reasonably is the key to ensuring the scientific and accuracy of the model. Under the limited information, we classify the factors that have an impact on the cost, which mainly be divided into the following aspects: engine performance indicators, magnitude indicators, and reliability indicators[2].

Table 1 – Cost impact parameter table

performance indicators	magnitude indicators,	reliability indicators
EP Engine power	Maximum diameter	Mean time between failures
Fuel consumption rate	Length	Air parking rate
Power-to-weight ratio	Height	Repair rate
Boost ratio	Weight	Unplanned reissue rate
Temperature inlet turbine	Frontal area	
Air flow		

The independent variables selected in model building need to meet the following conditions: 1. Large correlation with dependent variables; 2. Small correlation with each other; 3. Easy to obtain; 4. Quantifiable. Based on the conditions above, we select one of the engine performance indicators, the magnitude indicators and the reliability indicators factor as the independent variable of the model during the modeling process.
For the reliability indicator part, the air parking rate is a unique reliability indicator of the engine system, which may be caused by its own or the accessory system failure. Because the air parking failure will have a serious impact on the safety of the aircraft, so it is the most important reliability parameter for both military and civilian engines [3]. The repair rate is a reliability indicator for the use cost of the engine. A lower repair rate can reduce the engine life cycle cost and the cost of airlines. Unplanned reissue rate, which refers to the number of reissues within the unplanned period. As for the replacement reason, it is not only caused by the failure of the engine itself, but also by some reasons of the aircraft's other systems or the aircraft itself. The mean time between failures refers to the length of time between two failures. It is the most widely used parameter, but again, the cause of the failure cannot be determined. Through the analysis above, we come to the conclusion that, the air parking rate are the best indicator that reflects the reliability of the aircraft engine itself. Therefore, we choose it as the reliability factor of the production cost model.
In terms of magnitude indicator, according to the experience of typical domestic and foreign production cost evaluation models, weight indicator is an important cost influencing factor. Therefore, in this model, we also choose weight as the magnitude factor of the cost model.
As for performance indicator choose work, we do it under the premise above. In the previous analysis, we chose the air parking rate and weight as the dependent variables of the cost model. It can be seen that the power-to-weight ratio has a greater correlation with weight, and it has been learned from experience that the fuel consumption rate and the total boost ratio have a small correlation with the air parking rate index [4]. According to the principle of independent variable selection, we select the fuel consumption rate index as the performance factor of the cost model.
Through the above ideas, we finally determined that the parameters of the production cost estimation model are the air parking rate, weight and fuel consumption rate.

3.2 Multiple linear regression analysis

Among the multiple current parametric models, the model proposed by RAND Corporation in 1982 is the most widely used, which has been maturely used by RAND Corporation in many cost estimation practice. Considering that the civil aviation engines is not many and the data that can be collected is limited, we learn from the modeling ideas proposed by RAND Corporation in 1982 and use multiple linear regression analysis to establish a civil aviation engine production cost estimation model.

Multiple linear regression analysis is a theoretically mature statistical prediction technology, which plays a pivotal role in the modeling research of engine cost. It has always been the most effective and widely used modeling method ever, and the cost models still in use are almost all multiple regression models.

1）The form of the multivariate linear model

The multivariate linear model is expressed in matrix form, as

$$\begin{cases} \mathbf{y} = \mathbf{Xb} + \mathbf{e} \\ E(\mathbf{e}) = \mathbf{0} \\ \operatorname{cov}(\mathbf{e}) = \psi \end{cases} \quad (1)$$

Among them, y is called the observation vector; matrix X is called the regression design matrix; e is called the observation error; b is called the parameter vector; ψis called the error-covariance matrix. We assume that e is an independent normal random vector, ， and is an unit matrix of order n, that is, the random error satisfies the Gauss-Markov assumption.

If the X column is full rank, the least square solution of formula (1) is (the derivation process is omitted)

$$\hat{\mathbf{b}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \quad (2)$$

According to the Gauss-Markov theorem, the least squares solution of (2), is an unbiased estimate of the minimum variance of b.

2）Model hypothesis testing

In the process of multiple linear regression modeling, model hypothesis testing is an indispensable work. Only the multiple linear regression model satisfies the hypothesis test conditions that is credible. Common test models include goodness-of-fit test, F-test and t-test.

A. goodness-of-fit test

The goodness of fit is defined as

$$R^2 = \frac{\mathrm{SSR}}{\mathrm{SST}} = 1 - \frac{\mathrm{SSE}}{\mathrm{SST}} \quad (3)$$

The goodness of fit is the percentage of the explainable variation in the total variation. SST is the sum of squares of total variation, and its degree of freedom is (n-1); SSR is the interpretable sum of squares of variation, and its degree of freedom is k; SSE is the residual sum of squares, and its degree of freedom is (n-k-1)

The goodness of fit coefficient is one of the parameters that measure the quality of the model. The closer to 1, the higher the fit of the regression model. The size of is related to the number of independent variables. The more of the independent variables, the larger of , but it does not mean that the quality of the model is higher. In the case of a certain sample size n, increasing the independent variable will inevitably reduce the degree of freedom, so that the statistical trend of the data is less likely to appear. Therefore, many statisticians advocate that using as few independent variables as possible in regression modeling.

In order to make the degree of freedom factor be considered in the goodness of fit coefficient, the adjustment degree of fit coefficient is defined

$$\bar{R}^2 = 1 - \frac{\mathrm{SSE}/(n-k-1)}{\mathrm{SST}/(n-1)} \quad (4)$$

When n is very large and k is very small, the difference between and is not very big; However, when n is small and k is large, is much less than. It is time to consider reducing or adjusting the variable.

B. F-test

The purpose of F test is to test whether the dependent variable y has a linear relationship with the independent variable. The null hypothesis and the opposite hypothesis are as follows.

$$\mathrm{H}_0: \ \mathbf{b} = \mathbf{0}; \qquad \mathrm{H}_1: \ \mathbf{b} \neq \mathbf{0} \qquad (5)$$

The test statistic is:

$$F = \frac{\mathrm{SSR}/k}{\mathrm{SSE}/(n-k-1)} \sim F(k, n-k-1) \qquad (6)$$

For a given significance level, if $F \le F_\alpha(k, n-k-1)$,then H_0 is accepted, that is, the dependent variable y has no significant linear relationship with the independent variable; On the contrary, it negates H_0 and considers that the dependent variable y has a linear relationship with the independent variable.

C. t-test

The purpose of t-test is to test whether the global parameter of each independent variable is significant zero, that is, to determine whether has significant explanatory ability to Y. The null hypothesis and the opposite hypothesis are as follows.

$$\mathrm{H}_0: \ \mathbf{b}_j = \mathbf{0}; \qquad \mathrm{H}_1: \ \mathbf{b}_j \neq \mathbf{0}$$

The test statistic is:

$$t = \frac{b_j}{S(b_j)} \sim t(n-k-1) \qquad (7)$$

For a given significance level, if $|t| \le t_{\frac{\alpha}{2}}(n-k-1)$, then accept H_0 , that is, the parameter $\mathbf{b}_j$ is significantly zero; On the contrary, it is negated H_0 that the parameter $\mathbf{b}_j$ is significantly non-zero.

4. Cost model building

According to the model construction idea, we selected the air parking rate, engine weight and fuel consumption rate as the model dependent variables, and use the multiple linear regression analysis method to establish the production cost estimation model.

4.1 Data collection

The production cost of an engine is the internal data of engine manufacturer enterprises, which is the basis for the engine manufacturer to conduct commercial competition and obtain market competitiveness. As a result, it is not available to obtain a large number of the production costs data of engines from open sources.

The price of the aero-engine is mainly affected by the cost of the engine. Therefore, to a certain extent, the price of the engine reflects the size of its cost. Because engine costs are hard to obtain, in order to meet the modeling requirements, we collected the engine price as the basic data in the modelling.

At the same time, after a long period of accumulation and multi-channel data collection for the air parking rate, engine fuel consumption rate and weight. We finally form a set of engine production costs and parameter data as follows [5]. For some reasons, we do not list the engine type in the table.

The data in this table are obtained through domestic authoritative channels and are recognized as true data. Measure unit of fuel consumption rate is kg/ (daN.h); Measure unit of weight is “kg”; air parking rate is of thousands rate; Engine price is in "USD"，2012.

Table 2 – Engine data sheet

NO.	Engine price（USD）	Air parking rate	Fuel consumption rate	Weight
1	8,500,000	0.0013	0.614	2384
2	15,000,000	0.002	0.279	5443

3	15,000,000	0.001	0.279	5642.3
4	8,500,000	0.001	0.585	2242
5	15,000,000	0.002	0.571	4748
6	14,000,000	0.004	0.345	5091
7	17,000,000	0.001	0.562	7825

4.2 Model building

In the process of establishing the model, it is necessary to go through continuous attempts and explorations, discover the rules in the explorations, and further revise and improve the modeling ideas, and finally achieve more modeling state.

In the initial exploration of this modeling, we used the data in Table 2 above directly. We take the Engine price data in the above table as the dependent variable, the Air Parking Rate, Fuel Consumption Rate and Weight as the dependent variables to conduct regression modeling. There are many regression problems on the results, but at the same time it gives us a lot of enlightenment. The modeling results are obtained as follows:

Regression statistics	
Multiple R	0.962630602
R Square	0.926657676
Adjusted R Square	0.853315352
standard error	1298124.496
observed value	7

Variance analysis					
	df	SS	MS	F	Significance F
regression analysis	3	6.38732E+13	2.12911E+13	12.63469201	0.032967555
residual	3	5.05538E+12	1.68513E+12		
total	6	6.89286E+13			

Regression result						
	Coefficients	standard error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	6078248	3203928.27	1.89712	0.1540678	-4118082.0	16274577.
X Variable 1	2.79E+08	529810664.	0.52726	0.6345118	-14067408	19654471
X Variable 2	-1874254	4061875.95	-0.46142	0.675892	-14800956.	11052448.
X Variable 3	1590.381	296.311065	5.36726	0.0126600	647.38667	2533.3747

From the results above, it can be seen that the regression results have the following shortcomings: 1. The total standard error and the standard error of each coefficient are too large; 2. The absolute values of coefficients and constants obtained by regression are too large. From the above results, we can infer that the measurement magnitude of Engine price, Air parking rate, Fuel consumption rate and Weight is different, direct fitting cannot eliminate the influence. On the other hand, the linear relationship between Engine price and Air parking rate, Fuel consumption rate, Weight is not obvious. Thus, we need to do some work to solve the problems.

We went through countless modeling explorations. We Constantly discover problems, discover rules, and optimize modeling and finally, we combine useful continuous exploration and experience summary and processed the parameters as follows before modeling：1. Actually, the main fault of the engine is the air stop rate. Therefore, we make the following assumptions and processing: Air parking rate = failure rate =λ, Reliability $R = e^{-\lambda}$, R/1-R as an independent variable. 2. take the log of both the independent variable and the dependent variable. The sorted data table is shown below. Y is the logarithmic value of Engine price, X1 is the logarithmic value of R/1-R, X2 is the logarithmic value of Fuel consumption rate, X3 is the logarithmic value of Weight.

Table 3 – Improved data sheet.

NO.	dependent variable Y	independent variable X1	independent variable X2	independent variable X3
1	6.929418926	2.885774326	-0.211831629	3.377306251
2	7.176091259	2.698535637	-0.554395797	3.735838334
3	7.176091259	2.999782835	-0.554395797	3.751456174
4	6.929418926	2.999782835	-0.232844134	3.350635608
5	7.176091259	2.698535637	-0.243363892	3.67651071
6	7.146128036	2.39707113	-0.462180905	3.706803097
7	7.230448921	2.999782835	-0.250263684	3.893484346

In the process of modeling, it is generally necessary to select part of the data as modeling data and part of the data as validation data. Due to the small amount of effective data that could be collected in this study, we chose the first six groups of data as modeling data and the seventh group of data as verification data.

The first six groups of data in the above table were fitted by multiple linear regression, and the fitting results were shown as follows:

$$Y=-0.0225X1+0.1770 X2+0.8046X3+4.1966$$

It means:

$$\log P=\log(R/1-R)^{-0.0225}+\log A^{0.1770}+\log W^{0.8046}+4.1966=\log(R/1-R)^{--0.0225}\times A^{0.1770}\times W^{0.8046}\times 10^{4.1966}$$

So, the production cost model of civil aircraft engine is

$$P= (R/1-R)^{-0.0225}\times A^{0.1770}\times W^{0.8046}\times 10^{4.1966}$$

P is the aero-engine price, in USD,2012; R is the reliability defined by Air parking rate when Failure rate =Air parking rate. A is fuel consumption rate, in kg/ (daN.h). W is the weight of the aero-engine, in kg.

4.3 Model test

According to the principle of multiple linear regression analysis, it can be seen that the multiple linear regression results can only be considered scientific and effective if they satisfy the goodness of fit test, F-test and T-test. Therefore, we test and analyze the detailed results of the above regression to ensure the credibility of the regression. Detailed results of the above model fitting are as follows：

Regression statistics	
Multiple R	0.996828
R Square	0.993667
Adjusted R Square	0.984166
standard error	0.01561
observed value	6

Variance analysis					
	df	SS	MS	F	Significance F
regression analysis	3	0.076463	0.025488	104.5941	0.009485
residual	2	0.000487	0.000244		
total	5	0.07695			

Regression result						
	Coefficients	standard error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	4.196603	0.306294	13.70122	0.005285	2.878726	5.514481
X Variable 1	0.022465	0.035919	0.625435	0.595539	-0.13208	0.177012
X Variable 2	0.176985	0.074279	2.382685	0.140065	-0.14261	0.496583
X Variable 3	0.804625	0.073565	10.93759	0.008256	0.4881	1.121149

According to the above regression results, the regression model is highly reliable, mainly manifested in the following aspects: R2 is close to 1, and the standard error value is small, indicating that the model fitting degree is very good; is close to , indicating that the number of independent variables is reasonable; The values of SS and MS are very small, indicating that there is no big deviation in the model; A small F value indicates a high confidence of the model; The standard error of regression coefficient of each independent variable is small, indicating that the degree of fitting between independent variable and dependent variable is high.

At the same time, we used the seventh group of data in Table 3 for model instance verification. The independent variables of the seventh group of data, which were obtained from truthful data of the Air parking rate, Fuel consumption rate and Weight deformation, were input into the above regression results. The estimated value of Y' is 7.3525, and the actual value of Y is 7.2304. The error is calculated as follows

$$E = \frac{Y' - Y}{Y} \times 100\% = 1.68\%$$

The value of 1.68% is small enough. Therefore, we can conclude the model had a high degree of fitting and high reliability.

5. Summary

In this paper, we take the domestic demand as the traction, based on the current domestic and foreign modeling concepts and methods, absorb the mature modeling experience, establish the production cost estimation model in the development stage. The model has the following advantages: 1. The model considers the reliability factors. This model can reflect the quantitative relationship between production cost and reliability and fill the research blank at home. It provides an effective decision-making basis for the trade-off between production cost and reliability and performance level in the development stage.

2. The model regression results with high reliability. This modeling process is based on sufficient real data, using multiple linear regression analysis. The regression test results are good, and the model is verified with real data, and the error is very small.

3. The production cost estimation model has high operability. The number of independent variables of the model is reasonable, and the value of independent variables can be obtained in the development stage, so the model has a strong practicability for the production cost estimation in the development stage.

6. Contact Author Email Address

The contact author is Xiaocui-Li, email address is majiangmian@126.com .

7. Copyright Statement

References

[1] Sun Mingxia, Liang Chunhua. Innovative engines under the VAATE program. Aeronautical Science and Technology, 05(2011):18-21.

[2] Ao liang, Zhou Shenghan. The relationship model between the development cost and reliability of civil aircraft.

[3] Ji binghan. Research on Reliability Data Analysis and Prediction Method of Civil Aviation Engines. Diss. Civil Aviation University of China.

[4] Zhang hong, Ji Binghan and Cui Dongze. Calculation method of reliability design index for civil aero-engine. *Science Technology and Engineering*, v.19;No.498.29(2019):80-85.

[5] Lin Zuoming. World Aero Engine Handbook. 4th edition, Publisher, 2012.