The Aeronautical Journal (2023), **127**, pp. 268–288
doi:10.1017/aer.2022.56

REGULAR PAPER

Parametric cost estimation models of civil aircraft for the preliminary aircraft design phase

O. Al-Shamma[1,*] and R. Ali[2]

[1]University of Information Technology and Communications, Baghdad, Iraq and [2]Coventry University, Coventry, UK
[*]Corresponding author. Email: o.al_shamma@uoitc.edu.iq

Received: 25 August 2021; **Revised:** 26 April 2022; **Accepted:** 27 April 2022

Keywords: Parametric cost analysis; Aircraft cost estimation; Preliminary aircraft design; Correlation analysis; Multiple linear regression

Abstract
In the preliminary design phase of aircraft design, estimating the production cost accurately is a challenging task. At this stage, many design parameters that affect the overall cost are still undefined. This paper establishes cost-estimation models for civil, commercial aircraft using a parametric cost analysis (PCA) approach. Aircraft are characterised based on their size, ranging from a wide body to executive jets, into four categories. Key design parameters, such as maximum take-off weight, number of passengers, range, wing area, span, fuselage length, to name a few, are likely to be available in the preliminary design stage and significantly impact the aircraft design. These variables either directly or indirectly affect the overall production cost or performance. The PCA approach includes both correlation and multiple linear regression techniques. The empirical models thus developed were able to predict the aircraft cost with an error of less than ±4% for all aircraft categories considered. Two aircraft in each defined category were not part of the PCA models and were used to verify the models. The proposed models provide the ability to estimate the aircraft cost quickly in the early stages of the preliminary design phase and provide the possibility of performing parametric studies involving the key variables to determine the cost sensitivity to the main design parameters.

Nomenclature

b	wingspan (m)
$C_{est.[L]}$	cost estimate using linear regression (M$)
$C_{est.[Corr]}$	cost estimate using correlation analysis (M$)
$C_{est.[p-value]}$	cost estimate using p-value analysis (M$)
$C_{est.x}[y]$	cost estimate where suffix x denotes Aircraft Type A, B, C, or D, and suffix y denotes Linear regression (L), correlation analysis (Corr.) or p-value analysis (p-value) (M$)
D_f	fuselage diameter (m)
F	fuel capacity (L)
L_f	aircraft length (m)
M_n	Mach number
PAX	number of passengers (single class)
R	range (km)
S	wing area (m^2)
T	engine thrust (kN)
W_{to}	maximum take-off weight (tonnes)
λ	term coefficient in the correlation model

1.0 Introduction

Developing a new aircraft is a complex process since different design configurations have different technological and financial implications. Estimating the overall aircraft cost at the preliminary design stage is challenging using the standard cost forecasting techniques due to the lack of available sampled data. However, for the potential success of any aircraft program, the cost is considered a crucial indicator. In the early stages of the project, the program evaluation team is unlikely to know the design details and hence manufacturing implications. Amidst such uncertainties, assessing the viability of a program is a tricky task, which needs to be predicted accurately and timely. Consequently, a reliable estimating technique that enables cost-forecast is of high importance for any new commercial aircraft development. Availability of accurate costs decreases the risks associated with the program investment and enhances knowledge and accuracy, facilitating decision-making tasks [1]. Much of the cost estimation techniques are available for military aircraft; however, cost prediction models do not exist for commercial aircraft. If they do, they exist in a very rudimentary form. Therefore, it is necessary to develop a cost estimating technique for civil aircraft, which could be reliable amidst scant information and limited data, as an aid to evaluating the viability of new programs, as a function of critical parameters likely to have the most significant effect [2].

Methods for estimating the product cost in the fast-moving consumer goods market and general manufacturing are well established, but costs are unsuitable for predicting civilian aircraft due to small-scale production and the shortage of sample data. Insufficient information could lead to a sizable forecasting error using the common regression technique. It should be noted that several design parameters affect aircraft performance and costs. Moreover, predicting the production cost throughout the program development is desirable, as the key variables often change to reflect altered performance requirements. As the design complexity increases, the capital investment required for a successful program increases. Accordingly, the aircraft design process must accommodate design decisions based on the cost data and control the effective cost parameters during the design process to decrease the development cost. At the beginning of the design process, obtaining the cost is meaningful, as more than 70% of the aircraft cost lifecycle is assigned through the conceptual design phase [3].

Several methods for estimating aircraft cost have been put forward during the recent past. Researchers categorised these techniques into several approaches. Tirovolis et al. [4] categorised the existing techniques into two approaches: detailed cost accounting techniques [5] and parametric methods optimised for minimal cost [6]. Castagne et al. [7] categorised them into three approaches: detailed, parametric and analogous. In the first (detailed) approach, the material quantities and prices, the labour and rates, are estimated to determine the estimated direct costs of the activity or the product. More specifically, this approach requires considerable detailed information about the product, which turns out to be very costly and time-consuming, but it is the most accurate estimate. When the required work is decomposed into elementary activities, operations, or tasks to estimate the product cost, it is called Activity Based Costing (ABC). For estimation purposes, this technique (ABC) is effectively employed in the design stage as it is commonly utilised as a part of the total cost management [8]. In the second (parametric) approach, the cost of a product is estimated using past data and trends used for forecasting purposes. The approach comprises estimating cost relationships developed and applied to obtain the relations between the measurable attributes/parameters, the schedule, and the cost [9]. In addition, this approach was further applied to estimate the design process cost [10]. When using this technique, the cost estimates can be obtained quickly if the estimating cost relations are available.

In contrast, a significant effort is required for gathering past data and generating relationships. For products using up-to-date technologies, this approach is not well suited. The third (analogous) approach analogises with a similar existing product to estimate the cost. It works very well with novel products, but it requires a complete familiarity and expert judgement with the product [11].

Recently, Xie et al. [2] categorised the cost estimating techniques into six approaches: simulation, equalisation engineering value rate, empirical estimation, analogy, engineering estimation and parametric estimation. The simulation approach can evaluate multiple aspects during the preliminary design phase. Its main shortcoming is that it requires expert experience in a variety of fields. The equalisation is applied in the feasibility demonstration stage having a template structure, but its outcomes are not well behaved for various reasons. The empirical estimation approach is fast and low-cost, mainly applied in the detailed design phase. One fundamental limitation is due to subjectivity [2]. However, analogy estimation is the simplest among these approaches, mainly when data are challenging to achieve or acquire. It is easy and less time-consuming to estimate the overall project cost. However, accuracy seems deficient due likelihood of two projects being the same [11].

Lastly, the parametric estimation approach employs a mathematical model to extract the correlation between aircraft characteristics and development costs. It is based on statistical analysis, which requires parameters with significant effects to be assigned. This is followed by setting up the model technical performance factors and the historical cost data. Although the neural network approaches achieve successful estimation, they can recognise the relationship between the parameters and the development cost, but it has a poor estimation accuracy with limited data. In contrast, the engineering estimation approach is suitable for the latter stages of the design process, where the data are sufficiently refined [12]. For instance, the neural network was used to solve cost estimation challenges in the early civil building design process [13]. The researchers noticed that the estimation accuracy improves when the number of samples increases.

Furthermore, the neural network was combined with the regression technique to obtain a successful model for urban railways cost estimation [14]. Unfortunately, the neural network has not successfully estimated the developing aircraft cost due to the shortage of information, the number of samples, and the difficulty of obtaining the relevant cost parameters. This paper presents the parametric cost analysis (PCA) to estimate the aircraft cost for various aircraft categories to overcome such issues. It is suited during the preliminary design phase. Both correlation and multiple linear regression techniques are employed in the PCA method to investigate the critical design parameters for establishing the cost models. A simple regression model, which is widely used as a first, easy, is also presented in this paper for comparison purposes and highlights the PCA model accuracy. A comparative summary of the presented techniques is also introduced.

It should be noted that the models developed in this paper for estimating the designed aircraft cost are based on adjusted prices for the year 2018. In general, all market measures, such as inflation/deflation, CPI, etc., directly affect the price of the new product. These measures can be factored appropriately after estimating the newly designed aircraft price. In a similar way, the new cost of new technologies that either result in a price increase or decrease can be accounted for by using an appropriate weighting factor.

2.0 Aircraft categories and design parameters

The first and foremost requirement to obtain accurate cost estimation models is to have current and accurate data. Since it is meaningless to estimate the cost of an old or retired aircraft as it has different characteristics (i.e. old technologies and different materials). Thus, all the aircraft used in this paper are not just in service, but also most are still in production.

The second consideration is that the sampled data should have similar design characteristics and features. Therefore, the sampled data, which are utilised for both developing and testing processes, are categorised into four categories based on size, as follows:

Large-size aircraft (category A): It includes only Airbus and Boeing aircraft. Characterised by long-range, the number of passengers is more than 250, two-aisles seat configuration, maximum take-off weight greater than 200,000kg, and total aircraft length greater than 50m.

Mid-size aircraft (category B): It includes mainly Airbus and Boeing aircraft. Characterised by medium-range, the number of passengers between 110 and 250, single-aisle seat configuration, maximum take-off weight greater than 50,000kg and less than 200,000kg, and total aircraft length in the range of 30 to 50m.

https://doi.org/10.1017/aer.2022.56 Published online by Cambridge University Press

Table 1. *List of aircraft used for establishing the cost-estimating models*

Category A (Large-size aircraft)	Category B (Mid-size aircraft)	Category C (Small-size aircraft)	Category D (Personal aircraft)
777-200ER	737-700	CRJ-700	Cirrus Vision SF50
777-300ER	737-800	CRJ-900	Cessna CJ1
777-8	737-900ER	CRJ-1000	Cessna CJ2+
777-9	737-MAX 8	E-170	Cessna CJ3+
787-8	737-MAX 9	E-175	HondaJet
A330-200	A319	ATR-42	Embraer Phenom 100
A330-800	A320	ATR-72	Embraer Phenom 300
A330-900	A321	An-158	Cessna CJ4
A350-900	E-190	M90	Gulfstream G100
A350-1000	E-195	M100	Cessna XLS+

Small-size aircraft (category C): It is characterised by short-range, number of passengers between 70 and 110, single-aisle seat configuration, maximum take-off weight less than 50,000kg, and total aircraft length between 20 to 40m.

Personal aircraft (category D): It is characterised by a mainly very short-range, number of passengers less than 15, maximum take-off weight of less than 10,000kg, and total aircraft length of less than 15m.

Table 1 lists the utilised aircraft for establishing the estimating aircraft-cost models segregated into the four aircraft categories A, B, C, and D.

There is a dearth of information in the early stages of the preliminary aircraft design, especially pertaining to cost estimation, which is affected by many design choices and design parameters. There are many aspects of aircraft design in the conceptual and preliminary design stages that directly impact the final cost. Most of the choices and design parameters are correlated to each other to different degrees. The multi-collinearity issue occurs when two or more parameters are highly correlated. This issue makes selecting the most important parameters difficult or causes problems ranking them based on their importance. Thus, several regression techniques assume that the dataset is free of multi-collinearity. In contrast, multi-collinearity does not significantly impact selecting the design parameters, as this work does not apply the parameter-ranking issue. More specifically, these parameters are derived from either customer requirements or competitor analysis or may even be mandated by legislation. In aircraft design, weight, geometry and performance are the most significant issues to be considered. Maximum take-off weight (W_{to}) and fuel capacity (F) are selected to account for the weight. Wing area (S), wingspan (b), aircraft length (L_f), fuselage diameter (D_f), and the number of passengers (PAX) are selected to account for geometry. Lastly, range (R), maximum cruise speed (Mach number, M_n), and engine thrust (T) are selected to account for performance aspects. These selected parameters, which are described in the following, significantly impact almost every aspect of design and cost. Besides, they are likely to be known/evaluated during the conceptual design phase and/or available during the preliminary design phase.

The maximum take-off weight (W_{to}, kg) is selected as the first parameter to be considered along with all design phases. The significance of this variable affects almost every aspect of aircraft performance, thrust loading, wing loading, take-off performance, climb and cruise performance, range, endurance, turning performance and other performance metrics. In addition, there are many methods for estimating the maximum take-off weight during the conceptual and preliminary design phases.

The number of passengers (PAX) is generally a variable available at the conceptual design phase as one of the critical requirements. The number of passengers determines, along with the fuselage diameter, determines the seating arrangement, and has a bearing on fuselage length. This variable has a pronounced effect on the aircraft cost, as the cost associated with the structure goes up. The indirect effect

https://doi.org/10.1017/aer.2022.56 Published online by Cambridge University Press

is the tail arm and its impact on dynamic and static stability, requiring larger tail volume coefficients, thereby increasing areas of the empennage and ultimately cost.

From an aircraft performance aspect, the aircraft range R (km) is also a design requirement when examining the Breguet Range Equation:

$$R = \left(\frac{V}{g}\right) \times \left(\frac{1}{SFC}\right) \times \left(\frac{L}{D}\right) \times \ln\left(\frac{W_0}{W_f}\right) \tag{1}$$

The range is affected by aerodynamics (flight speed (V) and gravity acceleration (g)), propulsion ($SFC =$ fuel mass-flow rate/thrust), lift (L) and drag (D), and structures (initial weight (W_0) and final weight (W_f)). This translates to efficient engines and better aerodynamics, requiring boundary layer control, lower drag, a more robust structure with lower weights to yield a higher range and therefore increased aircraft cost.

Maximum cruise speed (Mach number, M_n) is selected, affecting the mission or stage time. Increasing this variable implies the availability of more excess power to accelerate the aircraft after overcoming the drag. Inevitably bigger engines with increased thrust are required, which implies more weight for the engines and the structure as the wing loading goes up. Maximum cruise speed depends not only on the thrust, but also on the drag, which needs to be reduced.

Engine thrust T (N) is also one variable available at the onset of the preliminary design stage. The thrust available limits the total drag that can be permitted, affecting the aerodynamic design, which also affects structural design to a large extent.

Aircraft length L_f (m) is one of the leading design parameters, which is dependent on the number of passengers and aisle configuration, with standard fore and after body requirements.

Fuselage diameter D_f (m), this variable is mainly dependent upon seating requirements.

The aircraft wing area S (m^2) depends on wing loading that comes about from constraint analysis, requiring many simultaneous constraints to be met.

Wingspan b (m) has roots in the choice of aspect ratio and wing loading.

Fuel capacity F (L) needs to be determined. It affects the maximum take-off weight, affecting the range and performance of the aircraft. Due to volumetric storage requirements, the fuel capacity also affects the geometry of the aircraft wing and/or fuselage design.

3.0 The proposed approach

As mentioned earlier, there is a lack of information on the present-day civil and military aircraft cost. The last updated available data are that for Airbus (for the year 2018) and Boeing (for the year 2019), which are extracted from their official websites [15, 16]. For establishing accurate cost-estimation models, the year 2018 is set as the price basis in this paper. Note that the Airbus and Boeing aircraft prices are for the year 2018. They are listed in categories A and B are achieved from Ref. 17. For categories C and D, the aircraft prices are extrapolated from different websites, where the data is modified or updated to the year 2018. As mentioned earlier, to obtain the estimated aircraft price at the year of manufacture, all market measures and new technologies (increment/decrement) are factored appropriately after estimating the designed aircraft price, using the models presented in various categories.

In this paper, Parametric Cost Analysis (PCA) technique establishes the cost-estimation models for all aircraft categories [18]. PCA is a dimensionality reduction technique specifically designed to reduce several design parameters into a smaller set with extremely high accuracy. The proposed methodology employs grey correlation in conjunction with the p-value analysis, similar to Chen et al. [18]. More specifically, two main improvements were made to enhance the cost-estimation accuracy effectively. The first improvement is using extended sample data and classifying them into four categories based on the aircraft size. The second improvement is replacing some design parameters (so-called cost-driven factors), which are more efficient and aligned with customer requirements. As a result, the values of the design parameters are easily determined in the early stage of the conceptual design. In addition, these

https://doi.org/10.1017/aer.2022.56 Published online by Cambridge University Press

Table 2. *List of current aircraft take-off weights and their prices (2018)*

Aircraft category	Aircraft type	Price (M$)	W_{to} (t)	Aircraft category	Aircraft type	Price (M$)	W_{to} (t)
A	777-200ER	295.2	297.55	C	CRJ-700	42.6	34
	777-300ER	361.5	351.53		CRJ-900	47.3	38.33
	777-8	394.9	351.8		CRJ-1000	50.5	41.64
	777-9	425.8	351.5		E-170	42	38.6
	787-8	239	227.93		E-175	46.7	40.37
	A330-200	238.5	242		ATR-42	23	45.9
	A330-800	259.9	251		ATR-72	27.7	30.48
	A330-900	296.4	251		An-158	40.4	18.6
	A350-900	317.4	280		M90	45.3	23
	A350-1000	366.5	316		M100	47.3	42.5
B	737-700	85.8	70.1	D	Cirrus Vision SF50	2.3	2.72
	737-800	102.2	79		Cessna CJ1	5.5	4.85
	737-900ER	108.4	85.1		Cessna CJ2+	7	5.67
	737 MAX 8	117.1	82.191		Cessna CJ3+	8.45	6.3
	737 MAX 9	124.1	88.314		HondaJet	5.3	4.85
	A319	92.3	75.5		Embraer Phenom 100	5.6	4.8
	A320	101	78		Embraer Phenom 300	9.45	8.15
	A321	118.3	93.5		Cessna CJ4	9.4	7.76
	E-190	52	51.8		Gulfstream G100	12.5	11.1
	E-195	54.7	52.3		Cessna XLS+	11.3	9.15

design parameters are strongly correlated with estimated aircraft cost, and any change in their values will result in changes in the estimated cost produced.

Moreover, each aircraft category results in very different prediction equations due to design complexities. Conventionally only maximum take-off weight was used in cost prediction; incorporating other design variables will significantly improve prediction accuracy and its ability to perform parametric trade-offs. The simple linear regression technique commonly used in terms of ($/kg) in many aircraft design textbooks is also presented here as a fast cost-estimation method and for comparison purposes.

3.1 Linear regression

Table 2 lists the current aircraft prices (as in the year 2018) with their relevant maximum take-off weights.

Simple linear regression is used to determine the best empirical model. The graphical representation of the data in Table 2 is shown in Fig. 1. The equation is determined to be:

$$C_{est.[L]} = 1.0908 * W_{to} + 4.3515 \tag{2}$$

$C_{est.[L]}$ is the cost estimate using linear regression.

This method is easy, fast, and commonly used as a primary estimating cost method in the conceptual design phase. Most modern aircraft design textbooks use this method to determine the aircraft price, mainly calculating the direct operating cost (DOC) [19, 20]. The main disadvantage of this simple method is its poor accuracy. It is acceptable in the conceptual design phase; however, in the advanced stages of design, many trade-offs are needed. Each design trade-off has a cost implication, and better cost estimation models are necessitated to perform parametric analysis due to changes in the key variables. The ten key design variables listed in section 2 are assumed to have a considerable effect on the overall cost. It shall be seen later that not all variables are required to predict the cost. Only the variables

https://doi.org/10.1017/aer.2022.56 Published online by Cambridge University Press

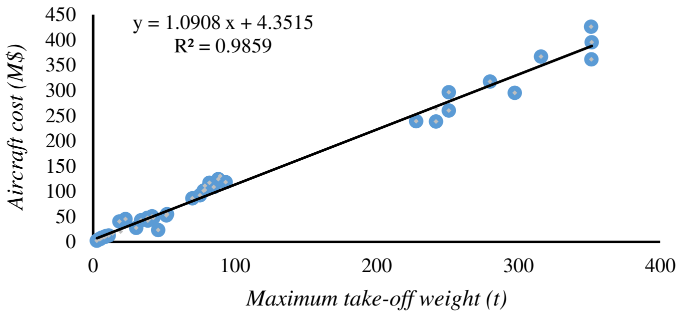

Figure 1. *Current aircraft prices versus their maximum take-off weights.*

that significantly impact the cost will be retained in the final models, since these variables are identified using correlation or p-value analysis.

3.2 Correlation analysis

Correlation analysis requires the correlation between system design parameters (independent variables) and the cost parameter (dependent variable) to be obtained and the degree of the correlation (ρ) between these design parameters. As mentioned in section 2, the ten design parameters for each aircraft category are utilised in this correlation analysis. The correlation degree has a value that varies between (-1) and ($+1$). The negative value (-1) means that the two parameters are entirely negatively correlated and vice versa. The range $|1| \geq \rho \geq |0.7|$ is indicative of a strong correlation between any two parameters, respectively [21]. This paper only includes strongly correlated design parameters to the cost parameter. Next, linear multivariable regression is applied using these strong parameters to obtain the coefficients of the correlation model. It is noted that if any two design parameters have a strong correlation between them, then these parameters have a multi-collinearity feature, which in turn, affects the accuracy of the estimated model assumed to be of the form:

$$C_{est.[Corr]} = \lambda_0 + \lambda_1 W_{to} + \lambda_2 R + \lambda_3 PAX + \lambda_4 L_f + \lambda_5 T + \lambda_6 M_n + \lambda_7 F + \lambda_8 S + \lambda_9 b + \lambda_{10} D_f \tag{3}$$

3.3 p-value analysis

The significance of every design parameter is determined via analysing the p-value of the regression model coefficients. The significance of the design parameter is high if it has a low p-value and vice versa. Note that eliminating low significant parameters will simplify the mathematical model. Analysing and determining the model using fewer parameters is much easier than the model of many parameters, as long as there is no appreciable decrease in the model accuracy.

However, this technique is an iterative procedure. The design parameter with the highest p-value is removed in each iteration if its value is above 0.05. The iteration procedure is stopped when all remaining design parameters in the current loop have p-values less than 0.05. These remaining parameters are utilised to establish the final regression model. For example, the first regression model comprises all ten design parameters. Due to the shortage of sample data in each category, the two design parameters with the lowest correlation degree are eliminated. Thus, the first regression model has only eight design parameters. Then, the second regression model includes only seven parameters after removing the one with the highest p-value. The third iteration will contain six, and so on. The procedure goes on until all parameters have p-values less than 0.05. The technique was applied to all aircraft categories in sequence. Each identified equation is of the form with one or more variables eliminated with p-values of >0.05.

https://doi.org/10.1017/aer.2022.56 Published online by Cambridge University Press

4.0 Aircraft cost-estimation models

Each category has ten aircraft (input sample data). Values of the input design parameters of each aircraft category are listed in Table 3 [17, 22]. These parameter values are employed to develop the estimated aircraft cost models using correlation analysis and multiple linear regression (p-value analysis). In addition, a simple regression model (Equation 2) is used to estimate the aircraft cost based on its maximum take-off weight for comparison purposes.

4.1 Correlation models

Based on the data presented in Table 3, the correlation analysis is applied first. For category A, the correlation matrix is shown in Table 4(a). The multi-collinearity feature is evident in the table, where several correlated coefficients between design parameters themselves have strong correlations ($\rho > 0.7$). These coefficients, presented in bold font in the table, show strong correlations, such as that noted between W_{to} and PAX.

Interestingly, the range variable has very little correlation with aircraft price. On the other hand, the strongly correlated design parameters to the price parameter, recognised by the bold and underlined font in the table, are used to establish the equation of the correlation model (category A). The final correlation model for aircraft category A is:

$$C_{est.A\,[corr]} = 2.294 * W_{to} + 0.1054 * PAX + 4.47 * L_f - 1.1 * T + 0.000075 * F - 1.448 * S + 0.17533 * b - 239.421 \quad (4)$$

A similar procedure was performed for the other aircraft categories. Tables 4(b), (c), and (d), show the correlation matrix for categories B, C and D, respectively.

Cost estimation Equations for Aircraft Categories B, C and D were determined as follows:

$$C_{est.B\,corr]} = 0.627 * W_{to} + 0.312 * PAX - 0.0736 * T + 4.404 * b - 0.263 * S - 0.00143 * F + 0.0067 * R - 5.827 * D_f - 100.436 \quad (5)$$

$$C_{est.C[corr]} = 1.28 * W_{to} - 2.946 * L_f - 0.00722 * F + 0.838 * PAX + 1.131 * T - 0.00646 * R - 0.6266 * S + 133.018 * M_n - 4.787 \quad (6)$$

$$C_{est.D[corr]} = 0.082 * b - 0.365 * PAX + 1.3 * L_f + 0.00152 * F - 0.016 * W_{to} + 0.008 * T + 1.324 * M_n - 0.00146 * R - 9.426 \quad (7)$$

4.2 Multivariable regression models based on p-value

The main objective of applying multiple linear regressing is to reduce the number of the design parameters in the corresponding correlation model with acceptable error. This reduction simplifies the final model with fewer variables in predicting the aircraft cost, analysing the design parameters and speeding up the decision-making task. As mentioned earlier, each category has ten aircraft (samples). Thus, the total number of dependent and independent variables should be less than the sample number to perform the linear regression correctly. Because there is only one dependent variable (the cost), only eight design parameters are included in the first cycle of the p-value analysis. Therefore, two design parameters with the lowest correlation degrees should initially be eliminated for each category. For example, in category A, the fuselage diameter D_f and the Mach number M_n are removed, while the aircraft length L and the Mach number M_n are removed in category B (see Table 4). The wingspan b and the fuselage diameter D_f are removed for category C, whereas the wing area S (insufficient data available) and the fuselage diameter D_f are removed in category D.

https://doi.org/10.1017/aer.2022.56 Published online by Cambridge University Press

Table 3. *The aircraft prices and the values of the input design parameters of all aircraft categories*

		Design parameters									
Aircraft category	Aircraft type	W_{to} (t)	R (km)	L_f (m)	D_f (m)	S (m^2)	b (m)	T (KN)	F (L)	PAX	M_n
A	777-200ER	297.55	13, 100	63.7	6.2	427.8	60.93	417	171, 171	440	0.89
	777-300ER	351.5	13, 600	73.9	6.2	436.8	64.8	513	181, 300	550	0.89
	777-8	351.8	13, 650	69.8	6.2	516.7	71.75	470	198, 000	535	0.89
	777-9	351.5	13, 500	76.7	6.2	516.7	71.75	470	198, 000	565	0.89
	787-8	227.93	13, 620	56.7	6.85	377	60.12	280	126, 206	359	0.9
	A330-200	242	13, 430	58.8	5.64	361.6	60.3	316	139, 090	406	0.89
	A330-300	242	11, 750	63.69	5.64	465	64	316	139, 090	440	0.86
	A330-800	251	15, 000	58.82	5.64	465	64	320	139, 090	406	0.86
	A330-900	251	13, 300	63.66	6	442	64.75	320	139, 090	440	0.86
	A350-1000	316	16, 100	73.88	6	464.3	64.75	432	159, 000	475	0.89
B	737-700	70.1	5, 500	33.6	3.8	124.6	35.8	89	26, 035	149	0.82
	737-800	79	5, 400	39.5	3.8	124.6	35.8	107	26, 035	189	0.82
	737-900ER	85.1	5, 500	42.1	3.8	124.6	35.8	120	29, 600	215	0.82
	737-MAX8	82.3	6, 570	39.52	3.8	127	35.9	123	25, 940	210	0.82
	737-MAX9	88.3	6, 570	42.16	3.8	127	35.9	123	25, 940	220	0.82
	A319	75.5	6, 800	33.84	4.05	124	35.8	120	29, 840	156	0.82
	A320	78	6, 100	37.57	4.05	124	35.8	120	27, 200	186	0.82
	A321	93.5	5, 920	44.5	4.05	128	35.8	147	30, 000	236	0.82
	E-190	51.8	4, 500	36.24	3	92.55	28.72	89	16, 420	114	0.82
	E-195	52.3	4, 260	38.65	3	92.55	28.72	89	16, 420	124	0.82
C	CRJ-700	34	2, 553	32.3	2.7	70.6	23.2	61.3	11, 200	78	0.82
	CRJ-900	38.33	2, 870	36.2	2.7	71.1	24.9	64.3	11, 200	90	0.82
	CRJ-1000	41.64	3, 000	39.1	2.7	77.4	26.2	64.3	11, 115	104	0.82
	E-170	38.6	3, 980	29.9	2.74	72.72	26	63	11, 760	78	0.82

Table 3. *Continued.*

Aircraft category	Aircraft type	Design parameters									
		W_{to} (t)	R (km)	L_f (m)	D_f (m)	S (m^2)	b (m)	T (KN)	F (L)	PAX	M_n
	E-175	40.37	4, 080	31.7	2.74	72.72	26	63	11, 760	88	0.82
	ATR-42	18.6	1, 326	22.67	2.9	54.5	24.57	40	5, 670	48	0.52
	ATR-72	23	1, 454	27.17	2.9	61	27.05	45	6, 330	72	0.49
	An-158	43.7	2, 500	30.8	3.1	87.3	28.56	67	16, 200	99	0.82
	M90	42.8	3, 770	35.8	2.95	85.3	29.2	78.2	12, 100	88	0.78
	M100	42	3, 540	34.5	2.95	85.8	27.8	78.2	12, 100	84	0.78
D	Cirrus Vision SF50	2.72	1, 100	9.42	1.65	–	11.8	8	1, 050	6	0.53
	Cessna CJ1	4.85	2, 800	13	1.6	22.3	14.4	8.74	1, 800	7	0.68
	Cessna CJ2+	5.67	3, 300	14.53	1.6	25	15.2	11.08	2, 220	9	0.72
	Cessna CJ3+	6.3	3, 780	15.6	2	27.32	16.26	12.5	2, 670	9	0.72
	HondaJet	4.85	2, 600	13	1.6	–	12.12	9.1	1, 400	7	0.73
	Embraer Phenon 100	4.8	2, 150	12.8	2	–	12.3	7.7	1, 450	7	0.71
	Embraer Phenom 300	8.15	3, 730	15.9	2.1	–	16.2	15.47	2, 800	8	0.78
	Cessna CJ4	7.76	4, 000	16.26	1.5	30.66	15.5	16.11	3, 300	10	0.78
	Gulfstream G100	11.1	5–470	16.9	1.55	29.4	16.64	18.9	5, 800	9	0.80
	Cessna XLS+	9.15	3–440	16	1.8	–	17.17	18.3	3, 800	9	0.77

Table 4. *The correlation matrices for the four aircraft categories*

(a) Category A

Design parameters	Price (M$)	M_w (t)	R (km)	*PAX*	L_f (m)	T (kN)	M_n	F (L)	b (m)	S (m^2)	D_f (m)
Price *(M$)*	1										
M_w (t)	**0.937**	1									
R (km)	0.340	0.277	1								
PAX	**0.940**	**0.939**	0.153	1							
L_f (m)	**0.954**	**0.915**	0.257	**0.922**	1						
T (kN)	**0.881**	**0.985**	0.211	**0.916**	**0.898**	1					
M_n	0.210	0.307	−0.013	0.098	0.239	0.305	1				
F (L)	**0.877**	**0.949**	0.147	**0.901**	**0.793**	**0.916**	0.272	1			
b (m)	**0.883**	**0.768**	0.433	**0.821**	**0.729**	0.661	−0.006	**0.781**	1		
S (m^2)	**0.818**	0.687	0.461	**0.741**	0.676	0.598	−0.284	0.689	**0.907**	1	
D_f (m)	0.139	0.184	−0.095	0.003	0.083	0.157	**0.729**	0.182	0.020	−0.117	1

(b)Category B

Design parameters	Price (M$)	W_{to} (t)	*PAX*	T (kN)	F (L)	R (km)	b (m)	S (m^2)	D_f (m)	L_f (m)	M_n
Price (M$)	1										
W_{to} (t)	**0.978**	1									
PAX	**0.957**	**0.964**	1								
T (kN)	**0.834**	**0.884**	**0.877**	1							
F (L)	**0.828**	**0.885**	**0.754**	**0.762**	1						
R (km)	**0.780**	**0.738**	0.607	0.698	**0.780**	1					
b (m)	**0.888**	**0.886**	**0.765**	0.657	**0.939**	**0.812**	1				
S (m^2)	**0.918**	**0.916**	**0.807**	**0.701**	**0.936**	**0.817**	**0.996**	1			
D_f (m)	**0.827**	**0.859**	**0.716**	**0.743**	**0.962**	**0.836**	**0.955**	**0.950**	1		
L_f (m)	0.574	0.605	**0.770**	0.651	0.246	0.058	0.198	0.262	0.159	1	
M_n	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1

(c) Category C

Design parameters	Price (M$)	W_{to} (t)	L_f (m)	F (L)	*PAX*	T (kN)	R (km)	S (m^2)	b (m)	D_f (m)	M_n
Price (M$)	1										
W_{to} (t)	**0.911**	1									
L_f (m)	**0.923**	**0.801**	1								
F (L)	**0.758**	**0.920**	0.591	1							
PAX	**0.842**	**0.876**	**0.847**	**0.798**	1						
T (kN)	**0.839**	**0.920**	**0.781**	**0.784**	**0.706**	1					
R (km)	**0.802**	**0.809**	0.589	0.647	0.535	0.783	1				
S (m^2)	**0.733**	**0.920**	0.687	**0.863**	**0.782**	**0.933**	0.638	1			
b (m)	0.157	0.459	0.223	0.354	0.414	0.518	0.282	0.670	1		
D_f (m)	−0.362	0.005	−0.307	0.115	−0.059	0.112	−0.250	0.348	**0.731**	1	
M_n	**0.915**	**0.904**	**0.743**	**0.895**	**0.755**	**0.788**	**0.779**	**0.726**	0.052	−0.296	1

https://doi.org/10.1017/aer.2022.56 Published online by Cambridge University Press

(d) Category D

Design parameters	Price (M$)	W_{to} (t)	R (km)	L_f (m)	T (kN)	F (L)	PAX	M_n.	b (m)	D_f (m)
Price (M$)	1									
W_{to} (t)	**0.984**	1								
R (km)	**0.918**	**0.913**	1							
L_f (m)	**0.951**	**0.904**	**0.931**	1						
T (kN)	**0.950**	**0.956**	**0.836**	**0.860**	1					
F (L)	**0.929**	**0.959**	**0.910**	**0.817**	**0.919**	1				
PAX	**0.815**	**0.737**	**0.806**	**0.889**	**0.778**	**0.712**	1			
M_n	**0.865**	**0.838**	**0.844**	**0.931**	**0.733**	**0.701**	**0.744**	1		
b (m)	**0.899**	**0.844**	**0.828**	**0.889**	**0.869**	**0.799**	**0.821**	0.694	1	
D_f (m)	0.280	0.236	0.143	0.281	0.172	0.104	0.035	0.257	0.269	1

Therefore, only eight parameters are included in the first round of p-value analysis. Next, the design parameter that has the highest p-value is removed. Then, the next cycle is performed with one less variable. The process continues until all remaining design parameters have a p-value less than 0.05.

Starting with aircraft category A, the first cycle showed that the PAX parameter has the highest p-value (0.858). In the second cycle, the wingspan b parameter was removed. Then, the fuel capacity F and the range R were removed in subsequent cycles. The reduced cost equation was determined to be:

$$C_{est.\,A\left[p-value\right]} = 1.92 * W_{to} + 4.9357 * L_f - 0.8883 * T + 0.15545 * S - 289.755 \tag{8}$$

By similar procedure, cost equations determined for the other aircraft categories are computed to be:

$$C_{est.\,B\left[p-value\right]} = 0.536 * W_{to} + 0.29 * PAX - 0.00145 * F + 0.00615 * R + 4.058 * b - 9.23 * D_f - 101.334 \tag{9}$$

$$C_{est.C\left[p-value\right]} = 1.28 * W_{to} - 2.946 * L_f - 0.00722 * F + 0.879 * PAX + 1.131 * T - 0.00646 * R - 0.6266 * S + 133.018 * M_n - 4.787 \tag{10}$$

$$C_{est.\,D\left[p-value\right]} = 1.42 * L_f - 0.00153 * R + 0.00156 * F - 0.37 * PAX - 8.833 \tag{11}$$

5.0 Results and discussion

Equations (4)–(11) are used to estimate aircraft costs which are compared with the actual aircraft costs to determine the difference (error) between them, and the error percentage of each aircraft is additionally calculated, as listed in Table 5.

Several measures and indexes are available in statistical analysis to evaluate the cost-estimation performance of such models. The first measure is the mean error (ME), which denotes the average of all differences between the estimated values and the actual values. The second measure is the mean absolute error (MAE), which uses the absolute difference values between actual and estimated costs. Mean percentage error (MPE) is another measure that refers to the average of the percentage differences between the estimated and the actual values. The actual differences contain positive and negative values, which can cancel each other out; as a result, they can be used as a biased measure in cost estimations. The fourth measure is the mean absolute percentage error (MAPE), commonly utilised in the model evaluation due to its simplicity and convenience. Finally, the R-squared is an additional index utilised to obtain how the

https://doi.org/10.1017/aer.2022.56 Published online by Cambridge University Press

Table 5. *The results of the three models for each aircraft category (Cost in M$)*

		$C_{est.\,[corr]}$			$C_{est.\,[p-val]}$			$_{s.\,[L]}$		
Aircraft type	Actual aircraft cost	Estimated cost	Error (Diff.)	Error (%)	Estimated cost	Error (Diff.)	Error (%)	Estimated cost	Error (Diff.)	Error (%)
777-200ER	295.2	289.65	−5.55	−1.88	292.17	−3.03	−1.03	328.39	33.19	11.24
777-300ER	361.5	359.97	−1.53	−0.42	361.94	0.44	0.12	387.04	25.54	7.07
777-8	394.9	390.85	−4.05	−1.03	393.03	−1.87	−0.47	387.33	−7.57	−1.92
777-9	425.8	424.17	−1.63	−0.38	426.51	0.71	0.17	387.01	−38.79	−9.11
787-8	239	236.32	−2.68	−1.12	237.61	−1.39	−0.58	252.74	13.74	5.75
A330-200	238.5	239.50	1.00	0.42	240.71	2.21	0.93	268.03	29.53	12.38
A330-800	259.9	264.12	4.22	1.62	266.96	7.06	2.72	277.81	17.91	6.89
A330-900	296.4	291.44	−4.96	−1.67	290.85	−5.55	−1.87	277.81	−18.59	−6.27
A350-900	317.4	309.65	−7.75	−2.44	314.03	−3.37	−1.06	309.32	−8.08	−2.55
A350-1000	366.5	367.58	1.08	0.30	370.54	4.04	1.10	348.43	−18.07	−4.93
737-700	85.8	85.83	0.03	0.03	86.04	0.24	0.28	81.24	−4.56	−5.32
737-800	102.2	101.89	−0.31	−0.30	101.88	−0.32	−0.31	90.91	−11.29	−11.05
737-900ER	108.4	108.44	0.04	0.04	108.20	−0.20	−0.19	97.54	−10.86	−10.02
737 MAX 8	117.1	117.05	−0.05	−0.04	117.46	0.36	0.31	94.38	−22.72	−19.40
737 MAX 9	124.1	124.01	−0.09	−0.07	123.67	−0.43	−0.35	101.03	−23.07	−18.59
A319	92.3	92.05	−0.25	−0.27	92.04	−0.26	−0.29	87.11	−5.19	−5.63
A320	101	101.10	0.10	0.09	100.78	−0.22	−0.22	89.82	−11.18	−11.07
A321	118.3	118.21	−0.09	−0.08	118.57	0.27	0.22	106.67	−11.63	−9.83
E-190	52	52.39	0.39	0.75	52.45	0.45	0.87	61.35	9.35	17.99
E-195	54.7	54.22	−0.48	−0.88	54.17	−0.53	−0.98	61.90	7.20	13.16
CRJ-700	42.6	42.61	0.01	0.03	42.61	0.01	0.03	42.01	−0.59	−1.38
CRJ-900	47.3	47.31	0.01	0.03	47.31	0.01	0.03	46.72	−0.58	−1.23
CRJ-1000	50.5	50.51	0.01	0.03	50.51	0.01	0.03	50.31	−0.19	−0.37
E-170	42	42.02	0.02	0.05	42.02	0.02	0.05	47.01	5.01	11.93
E-175	46.7	46.72	0.02	0.04	46.72	0.02	0.04	48.93	2.23	4.79
ATR-42	23	23.01	0.01	0.03	23.01	0.01	0.03	54.94	31.94	138.89

Table 5. *Continued.*

Aircraft type	Actual aircraft cost	$C_{est.\,[corr]}$			$C_{est.\,[p-val]}$			$s.\,[L]$		
		Estimated cost	Error (Diff.)	Error (%)	Estimated cost	Error (Diff.)	Error (%)	Estimated cost	Error (Diff.)	Error (%)
ATR-72	27.7	27.71	0.01	0.03	27.71	0.01	0.03	38.19	10.49	37.86
An-158	40.4	40.41	0.01	0.03	40.41	0.01	0.03	25.28	−15.12	−37.43
M90	45.3	45.32	0.02	0.04	45.32	0.02	0.04	30.06	−15.24	−33.64
M100	46.3	46.32	0.02	0.04	46.32	0.02	0.04	51.25	3.95	8.35
Cirrus Vision SF50	2.3	2.31	0.01	0.43	2.28	−0.02	−0.94	8.02	5.72	248.89
Cessna CJ1	5.5	5.64	0.14	2.55	5.56	0.06	1.11	10.34	4.84	87.98
Cessena CJ2+	7	6.93	−0.07	−0.97	6.88	−0.12	−1.66	11.23	4.23	60.43
Cessna CJ3+	8.45	8.39	−0.06	−0.66	8.37	−0.08	−0.94	11.91	3.46	41.00
HondaJet	5.3	5.21	−0.09	−1.76	5.24	−0.06	−1.08	10.34	5.04	95.07
Phenom 100	5.6	5.66	0.06	1.03	5.73	0.13	2.24	10.28	4.68	83.65
Phenom 300	9.45	9.49	0.04	0.41	9.45	0.00	−0.04	13.92	4.47	47.35
cessna CJ4	9.4	9.55	0.15	1.56	9.58	0.18	1.96	13.50	4.10	43.63
Gulfstream G100	12.5	12.49	−0.01	−0.11	12.51	0.01	0.11	17.13	4.63	37.04
Cessna XLS+	11.3	11.27	−0.03	−0.30	11.22	−0.08	−0.69	15.01	3.71	32.84

Table 6. *The measures' results of the three models*

		Aircraft category			
Model	Metric	A	B	C	D
Correlation	ME	−2.19	−0.0723	0.01	0.01
	MAE	2.8125	0.1728	0.0033	0.0664
	MPE	−0.66	−0.073	0.03	0.22
	MAPE	0.9672	0.2478	0.0051	0.9777
	R^2	0.997	0.9999	1	0.9993
p-value regression	ME	−0.08	−0.066	0.01	0.00
	MAE	2.9663	0.3163	0.0033	0.0746
	MPE	0.00	−0.065	0.03	0.01
	MAPE	1.0048	0.3881	0.0051	1.0781
	R^2	0.9967	0.9998	1	0.9990
Simple regression	ME	2.88	−8.4	2.19	4.49
	MAE	21.1	8.076	8.5341	0.4936
	MPE	1, 85	−5.98	12.78	77.79
	MAPE	6.81	8.8217	30.24	40.888
	R^2	0.9859	0.9859	0.9859	0.9859

linear model fits the values of the parameters (or set of observations). The metrics of the models under consideration for all categories are listed in Table 6.

Examining Table 6 shows that the correlation model produces the best prediction, which achieves lower MAE and MAPE in the four categories. In contrast, the *p*-value model shows slightly better results in both ME and MPE for categories A and D. In addition, the correlation model achieved the highest R^2 in all categories. Conversely, the linear regression model is the worst in all measures. Understandably, the correlation model has the highest number of design parameters in establishing its model equations. In contrast, the *p*-value regression findings show minimal differences compared to the correlation findings. It produced sufficiently acceptable results for use in preliminary design. Thus, either the correlation or the *p*-value regression model can be used, depending on the availability of required parameters.

The percentage error values in Table 5 are considered to highlight the model's accuracy in detail for all categories. For obtaining the overall error accuracies of the three models, the maximum and minimum percentage error values are extracted for all categories. The error accuracy is determined as the highest value between the Max and Min values. The variation range between the Max and Min values are listed, as well. Table 7 summarises these key statistical findings.

From Table 7, it is evident that the correlation regression models have the best error accuracies, typically less than ±3% for all aircraft categories. In contrast, the *p*-value regression models are also suitable for less than ±3% error. The difference between the two models is minimal, and it seems to be very close to each other. Therefore, both methods can be considered for early aircraft cost estimation. The *p*-value model is recommended since it has a lower number of design parameters. This finding certifies our earlier decision.

Figures 2 and 3 present the actual cost of the aircraft and the predicted cost using the models presented for aircraft categories A and B, respectively. Examination of these figures reveals that our cost-estimation models (correlation and *p*-value) are helpful during the preliminary design phase with minimal errors.

Table 7. *Error accuracies of the PCA model for all categories*

	Correlation model				*p*-value regression model				Simple regression model			
Air. Cate	Min error (%)	Max error (%)	Var. range (%)	Error accuracy (%)	Min error (%)	Max error (%)	Var. range range (%)	Error accuracy (%)	Min error (%)	Max error (%)	Var. range (%)	Error accuracy (%)
A	−2.44	+1.62	4.06	±2.5	−1.87	+2.72	4.59	±2.8	−9.11	+12.38	21.49	±12.4
B	−0.88	+0.75	1.63	±0.9	−0.98	+0.87	1.85	±1	−19.4	+17.99	37.39	±19.5
C	0.03	+0.05	0.02	±0.1	0.03	+0.05	0.02	±0.1	−33.64	+138.9	172.5	±139
D	−1.76	+2.55	4.31	±2.6	−1.66	+2.24	3.9	±2.3	+32.84	+248.9	215.9	±249

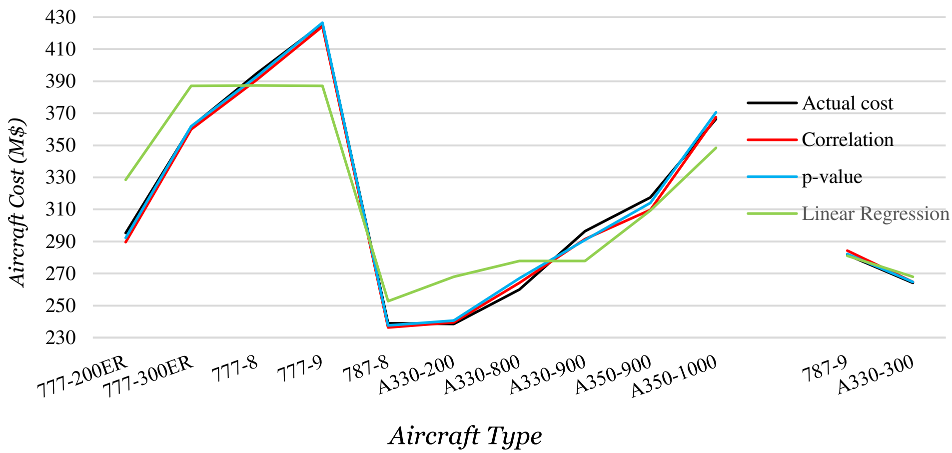

Figure 2. *Actual and Estimated (Correlation/p-value/Linear regression) Aircraft Cost – Category A.*

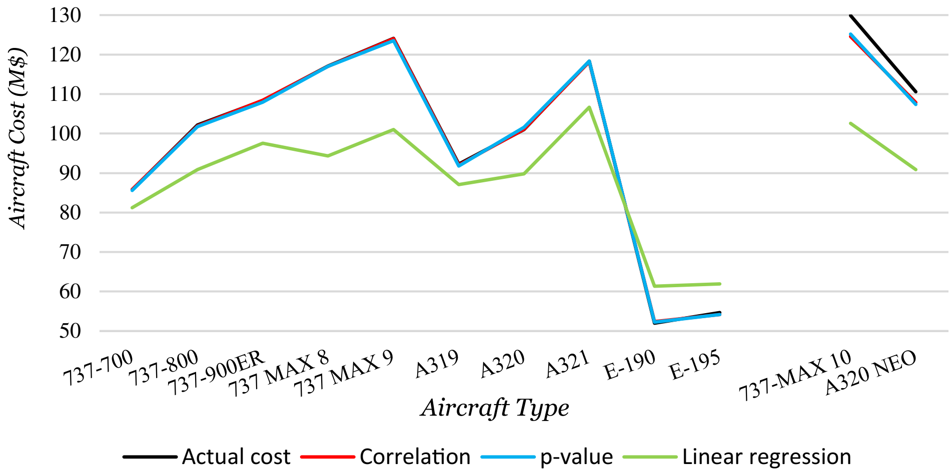

Figure 3. *Actual and Estimated (Correlation/p-value/Linear regression) Aircraft Cost – Category B.*

The models presented constitute a significant advancement over the existing models used in cost estimation. As mentioned earlier, the linear regression model used in the past is presented here just for comparison purposes, mainly highlighting its inherent inaccuracies.

6.0 Case Studies

Confirming the suitability of the developed models, some sample data (aircraft data) other than that used in developing the estimated aircraft cost model are considered for testing. These sample data, which include two aircraft for each category with their design parameter values, are listed in Table 8. Using the data in Table 8, equation (2), and (4)–(11), estimated aircraft costs are calculated (as shown on the right-side of Figs 2 and 3 for categories A and B, respectively), subtracted from the actual aircraft costs to determine the difference (error), and the error accuracy percentage of each aircraft for the three models are calculated, as presented in Table 9.

https://doi.org/10.1017/aer.2022.56 Published online by Cambridge University Press

Table 8. *The test aircraft (prices and design parameters)*

			Design parameters									
Aircraft category	Aircraft type	Actual aircraft price (M$)	W_{to} (t)	R (km)	L_f (m)	T (kN)	F (L)	PAX	M_n	b (m)	S (m^2)	D_f (m)
A	787-9	281.6	254	14, 140	62.8	320	126, 372	406	0.9	60.12	377	5.85
	A330-300	264.2	242	11, 750	63.67	316	139, 090	440	0.86	60.3	361.6	5.64
B	737-MAX10	129.9	89.77	6, 100	43.8	130	25, 800	230	0.82	35.9	127	3.8
	A320NEO	110.6	79	6, 500	37.57	120.6	26, 730	195	0.82	35.8	124	4.05
C	E175-E2	48.2	44.8	3, 700	32.4	67	11, 100	88	0.82	31	103	3.2
	SSJ 100-95	52	49.45	3, 540	29.9	71.6	15, 800	108	0.81	27.8	83.8	3.46
D	Cessna Latitude	16.75	13.97	5, 000	19	26.3	6, 435	9	0.78	22	50.4	2.2
	Embraer Praetor 600	20.4	19.44	7, 440	20.74	31.3	9, 700	12	0.83	21.5	44.85	2.3

https://doi.org/10.1017/aer.2022.56 Published online by Cambridge University Press

Table 9. *The error accuracies of the aircraft test cases*

			Correlation model			*p*-value regression model			Simple regression		
Cat.	Aircraft type	Actual aircraft cost	Estimated cost	Error (Diff.)	Error (%)	Estimated cost	Error (Diff.)	Error (%)	Estimated cost	Error (Diff.)	Error (%)
A	787-9	281.6	284.33	2.73	0.97	282.24	0.64	0.23	281.08	−0.52	−0.19
	A330-300	264.2	264.76	0.56	0.21	264.65	0.45	0.17	268.03	3.83	1.45
B	737-MAX 10	129.9	124.58	−4.32	−3.35	124.69	−4.21	−3.27	102.61	−27.29	−21
	A320 NEO	110.6	107.88	−2.72	−2.46	107.13	−3.47	−3.14	90.91	−19.69	−17.8
C	E 175-E2	48.2	47.12	−1.08	−2.23	47.12	−1.08	−2.23	53.75	5.55	11.51
	SSJ 100-95	52	53.49	1.49	2.87	53.49	1.49	2.87	58.8	6.8	13.08
D	Cessna Latitude	16.75	17.29	0.54	3.25	17.21	0.46	2.72	20.25	3.5	20.89
	Embraer Praetor 600	20.4	19.84	−0.56	−2.75	19.94	−0.46	−2.25	26.19	5.79	28.39

The findings from the aircraft used for model validation indicate that p-value regression is better than the correlation model for categories A, C and D, with an error of less than ±3%. For category B, the correlation model is slightly better, but both models give an error of less than 4%. In general, most textbooks state that an error of less than 10% is acceptable and a cost prediction error of less than 5% is significantly good, considering that not much detail is available at the early design stages.

6.1 Model sensitivity

Investigating the sensitivity of the established models requires examining each mathematical model. The linear feature of the models simplifies the computation of the estimating cost and allows performing parametric studies of the design parameter changes on the estimated cost. The degree of each parameter impact is based on its coefficient. The highest coefficient value has the highest impact on the estimated cost, and vice versa. In addition, the positive coefficient sign means that the parameter is directly proportional to the estimated cost, while the negative coefficient sign means inversely proportional to the estimated cost. However, starting with category A, the correlation and the p-value models share the same four design parameters. The W_{to}, L_f, and S are directly proportional to the estimated cost, while the T parameter is inversely proportional to the estimating cost.

Moreover, the L_f parameter has the highest coefficient value; therefore, it holds the highest impact. Increasing/decreasing the length by one meter means that the estimated cost will increase/decrease by 4.9M$, while a weight increment/decrement of one ton will cause a 1.9M$ increment/decrement in the estimated cost. The wing area has the lowest impact on the estimated cost.

In contrast, the W to has the highest coefficient (0.61) when considering category B. Notice that the PAX parameter has a coefficient value of around 0.3, but it holds the highest impact. Each additional passenger will increase the aircraft cost by one-third million dollars. In other words, the seat configuration of this category has six seats in each row (i.e. the cabin length and, in turn, the aircraft length will increase by around one meter); thus, the estimated cost will increase by around 2M$.

Category C shows that L_f has the highest coefficient but with a negative sign. Therefore, a one-meter increase/decrease in aircraft length will decrease/increase the estimated cost by around three million dollars. In contrast, the PAX variable has a positive sign with a value of around 0.88. More specifically, increasing one meter in length will increase the number of passengers by four (one row). Hence, an increase in cost by around 0.5M$ is the net impact, while the W_{to} will have the highest impact. Lastly, category D shows the same scenario as category C concerning L_f and PAX variables. Besides, improving the engine consumption will increase the range with no additional fuel and reduce the estimated cost. For more accurate cost estimation, the parameters W_{to}, T, S and M_n are additionally included in the correlation model. Thus, we can conclude that the maximum take-off weight has a pronounced parameter impact on the estimated aircraft cost, and it should always be closely monitored during the whole design process.

7.0 Conclusions

This paper presented the empirical models for estimating the cost of civil aircraft using the parametric cost analysis, where both correlation and p-value regression techniques were performed. The sample data were categorised into four categories for obtaining accurate models based on aircraft size. In addition, ten design parameters were considered that have a significant impact on the aircraft design and, therefore, the eventual cost. Finally, the correlation and the p-value analyses were applied to determine the significant design parameters that affect the cost. Among the ten identified design parameters, the maximum take-off weight has the highest impact on most cost-estimation models. As a result, the developed models can predict cost-estimation at a preliminary design phase with an overall error better than ±4%, which is significantly better than the available methods and provides an ability to perform parametric studies involving key design variables at the preliminary aircraft design stage.

References

[1] Chen, X., Huang, J., Yi, M. and Pan, Y. Prediction of the development cost of general aviation aircraft, *Aircraft Eng. Aerospace Technol.*, 2019, **91**, (4), pp 567–574. DOI: 10.1108/AEAT-09-2018-0248

[2] Xie, N.-m., Yin, S.-M. and Hu, C.-Z. Estimating a civil aircraft's development cost with a GM (1, N) model and an MLP neural network, *Grey Syst. Theory Appl.*, 2017, **7**, (1), pp 2–18. DOI: 10.1108/GS-11-2016-0049

[3] Asiedu, Y. and Gu, P. Product life cycle cost analysis: State of the art review, *Int. J. Prod. Res.*, 1998, **36**, (4), pp 883–908. DOI: 10.1080/002075498193444

[4] Tirovolis, N.L. and Serghides, V.C. Unit cost estimation methodology for commercial aircraft, *J. Aircraft*, 2005, **42**, (6), pp 1377–1386. DOI: 10.2514/1.12491

[5] Wu, H., Liu, Y., Ding, Y. and Liu, J. Methods to reduce direct maintenance costs for commercial aircraft, *Aircraft Eng. Aerospace Technol.*, 2004, **76**, (1), pp 15–18. DOI: 10.1108/00022660410514964

[6] Ross, T.E. and Crossley, W.A. Method to assess commercial aircraft technologies, *J. Aircraft*, 2000, **37**, (37), pp 570–579. DOI: 10.2514/2.2668

[7] Castagne, S., Curran, R., Rothwell, A., Price, M., Benard, E. and Raghunathan, S. A generic tool for cost estimating in aircraft design, *Res. Eng. Des.*, 2008, **18**, (4), pp 149–162. DOI: 10.1007/s00163-007-0042-x

[8] Feng, S. and Song, E. A manufacturing process information model for design and process planning integration, *J. Manuf. Syst.*, 2003, **22**, (1), pp 1–15. DOI: 10.1016/S0278-6125(03)90001-X

[9] Dean, E.B. Parametric cost deployment, Proceedings of the 7th symposium on quality function deploymen*t*, MI, USA, November 1995, pp 27–34.

[10] Roy, R., Bendall, D., Taylor, J., Jones, P., Madariaga, A., Crossland, J., Hamel, J. and Taylor, I. Development of airframe engineering CER's for aerostructures, Proceedings of the Second World Manufacturing Congress (WMC'99), Durham (UK), 1999, pp 838–844. DOI: 10.1007/978-1-4612-2936-0_12

[11] Dunk, A.S. Product life cycle cost analysis: The impact of customer profiling, competitive advantage, and quality of IS information, *Manag. Account. Res.*, 2004, **15**, (4), pp 401–414. DOI: 10.1016/j.mar.2004.04.001

[12] Cai, W. and Fang, W. Combination forecasting method for the development cost of aircraft, *Syst. Eng. Electron.*, 2014, **36**, (8), pp 1573–1579. DOI: 10.3969/j.issn.1001-506X.2014.08.20

[13] Ambrule, V.R. and Bhirud, A.N. Use of artificial neural network for pre-design cost estimation of building projects, *Int. J. Recent Innovation Trends Comput. Commun.*, 2017, **5**, (2), pp 173–176. DOI: 10.6106/jcepm.2014.4.4.009

[14] Mahalakshmi, G. and Rajasekaran, C. Early cost estimation of highway projects in India using artificial neural network, in *Sustainable Construction and Building Materials*, Springer, 2019, Singapore, vol. **25**, pp 659–672. DOI: 10.1007/978-981-13-3317-0_59

[15] https://www.airbus.com/newsroom/press-releases/en/2018/01/airbus-2018-price-list-press-release.html

[16] https://www.boeing.com/company/about-bca/

[17] https://jetlinemarvel.net/2018/01/25/airbus-boeing-average-list-prices-2018/

[18] Chen, X., Huang, J. and Yi, M. Development cost prediction of general aviation aircraft projects with parametric modelling, *Chin. J. Aeronaut.*, 2019, **32**, (6), pp 1465–1471. DOI: 10.1016/j.cja.2019.03.024

[19] Raymer, D. *Aircraft Design: A Conceptual Approach*, 6th ed, AIAA, 2018, Reston, USA. DOI: 10.2514/4.104909

[20] Kundu, A.K., Price, M.A. and Riordan, D. *Conceptual Aircraft Design: An Industrial Approach*, 1st ed, John Wiley & Sons, 2019, Hoboken, NJ.

[21] Chen, X., Yi, M. and Huang, J. Application of a PCA-ANN based cost prediction model for general aviation aircraft, *IEEE Access*, 2020, **8**, pp 130124–130135. DOI: 10.1109/ACCESS.2020.3008442

[22] Al-Shamma, O. Development of interactive aircraft design software for use in problem-based learning. PhD dissertation, University of Hertfordshire, UK, 2013. DOI: 10.18745/th.12108

Cite this article: Al-Shamma O. and Ali R. (2023). Parametric cost estimation models of civil aircraft for the preliminary aircraft design phase. *The Aeronautical Journal*, **127**, 268–288. https://doi.org/10.1017/aer.2022.56