



A cost estimation approach for aircraft design enhancement

Raffaele Abbate¹ · Maria Antonietta Turino¹ · Llewellyn Morse² · Marcello Fera¹ · Vincenzo Mallardo² · Roberto Macchiarelli¹

Received: 2 December 2022 / Accepted: 29 May 2023 / Published online: 7 July 2023
© The Author(s) 2023

Abstract

This paper presents a novel reliability-based parametric methodology for quantifying the reliability of cost estimates for new composite aircraft components. In recent years, the aircraft production sector has increased its attention on optimizing their composite manufacturing operations. It has become clear that a key factor in the success of these operations is the consideration that not only technical factors, but also economic ones as well are relevant. Composite manufacturing variability is greatly influenced by many of these factors, and since manufacturing with composites is significantly more difficult than with more traditional materials, there are many sources of uncertainty that could influence the reliability of manufacturing cost estimates for new composite aircraft components. Therefore, it is worth considering these sources of uncertainty during the cost estimation process and to quantify the reliability of the cost estimates. To demonstrate the proposed methodology, a numerical example featuring a real-life composite aircraft component from a Boeing 787, with real-life data, is presented. Results show that the proposed methodology can quantify the uncertainty associated with cost estimates for new composite aircraft components in an effective manner, thereby supporting engineers in optimising the cost of their designs, helping them avoid errors in budget definition, and enabling them to allocate resources more efficiently.

Keywords Parametric cost estimation · Composites · Design for manufacturing · Manufacturing cost · Optimisation

Abbreviations

CI	Confidence interval	y_i	Actual cost of the i 'th historical component
CERs	Cost estimation relationships	\hat{y}_i	Predicted cost of the i 'th historical component
CDF	Cumulative distribution function	β_0	Intercept of the regression model
DFM	Design for manufacturing	β	Coefficient of the regression model
LCC	Life cycle cost	ε_i	Term accounting for random error
MAPE	Mean absolute percentage error	k	Initial number of cost drivers
MSE	Mean square error	p	Final number of cost drivers
PCA	Principal component analysis	n	Number of observations included into the historical database

Symbols

\mathbf{Y}	Vector of the recorded manufacturing costs	R^2	Determination coefficient of the regression model
\mathbf{X}	Matrix of the cost drivers	R^2_{adj}	Adjusted determination coefficient of the regression model
		$H_{j,0}, H_{j,1}$	Null hypothesis and alternate hypothesis for the partial F test
		H_0, H_1	Null hypothesis and alternate hypothesis for the overall F test
		$t_{\alpha/2, n-(k+1)}$	Student's t-distribution parameter
		ρ	Pearson correlation coefficient
		λ	Threshold value for Pearson coefficient

✉ Maria Antonietta Turino
mariaantonietta.turino@unicampania.it

¹ Department of Engineering, University of Campania "Luigi Vanvitelli", Via Roma 29, 81031 Aversa, Italy

² Department of Architecture, University of Ferrara, Via Quartieri 8, 44121 Ferrara, Italy

\mathbf{x}_{new}	Vector of the cost driver of the new component
\hat{y}_{new}	Predicted cost of the new component
F_{critical}	Critical value of the Fisher distribution
df_1, df_2	Degrees of freedom of the Fisher distribution
α	Level of significance in the Fisher distribution
R_{required}	Desired reliability value for the cost estimation
C_{max}	Maximum budget
μ	Mean value of normal distribution
σ	Standard deviation of normal distribution
\mathbf{d}	Vector of the design variables of product
\mathbf{d}^U	Upper limit value for the design variables
\mathbf{d}^L	Lower limit value for the design variables
$\text{Cost}_{\text{manufacturing}}$	Manufacturing cost
C_{N_plies}	Regression coefficient related to the number of composite plies
C_{length}	Regression coefficient related to the length of the component
C_{radius}	Regression coefficient related to the corner radius of the component
C_{width}	Regression coefficient related to the width of the component
$C_{\text{flange_width}}$	Regression coefficient related to the flange width of the component
C_0	Intercept of the regression model

1 Introduction

A significant percentage of the total life cycle cost (LCC) of an aircraft is associated with the production of its components; studies have shown that this percentage is around 32% [1]. As demonstrated in [2], more than 70% of the production cost is determined in the design phase. Authors in [3] identified data and information related to several Design for Manufacturing (DFM) factors by analysing 100 case studies. These factors were then evaluated by three criteria concerning the importance according to experts, the applicability in operations and the percentage of cost reduction. According to [4], it makes much more sense to carry out a more comprehensive and accurate cost analysis during the early design phase to reduce costs before the start of the production phase. As proved by the authors in [5] the application of their algorithm allowed a production cost reduction of 18% by optimizing the product design. Moreover, a careful cost analysis in the design phase could be a point of agreement between designers and manufacturers as proved by the authors in [6] where

they investigated the influence on production cost of the product tolerances on which designers and manufacturers often disagree. Indeed, a reliable estimate of costs is one of the key principles to maximise a company's efficiency and, consequently, its profitability. This last aspect is highlighted in [7] where the authors investigated the interaction between the economic and manufacturing fields, highlighting the importance of having an estimate of the cost as soon as possible to submit a proposal to a competitive bidding process.

With the advancement of the aerospace industry and increasingly fierce competition in the market, presenting a design with a low production cost can have a significant impact on the market [8]. In recent years, composites have become more widely used in the aviation industry due to their inherent characteristics, i.e., good corrosion and fatigue resistance, higher specific strength, resulting in reduced aircraft weight and improved performance [9]. Indeed, according to [10] and [11] about 35% of metallic components of aircraft were replaced with composite ones in the latest models of aircraft. Although composite materials have proven their superiority in terms of weight, mechanical properties, and can provide significant savings in terms of fuel costs during the service life of the aircraft [12], their use is often limited due to their relatively high cost. In fact, they are about three times more expensive per kg than aluminium alloys in terms of raw material cost [13, 14]. Thus, cost estimation in the early design phase of a new aircraft is particularly important when composites materials are used [15, 16]. However, due to the greater complexity of the techniques and an application that is still in its early phases in the aeronautical field, the knowledge and information available for the cost estimation of composites is particularly scarce [4]. It is common practice to conduct a cost analysis to choose the best compromise between cost and performance [17, 18]. Aircraft development and production is considered a complex engineering procedure, characterised by many interconnected design steps with long and manual production cycles, high occurrence of technical difficulties, operational uncertainty factors, etc. The modern aircraft industry started to use cost estimation as a decision-making tool for processes such as production and maintenance, as well as structural analysis.

The European project MASCOT [19] (Modular multi-level cost Analysis Software for COmposiTe smart fuselage) aimed to contribute to this topic by developing a multi-disciplinary optimisation open-source software to estimate the cost of manufacturing activities of a transport composite fuselage during its design phase, integrating it in a wider design vision that involves also CAD representation, structural analysis, Structural and Health Monitoring (SHM) technologies and statistical confidence. Indeed, the design process of a fuselage is an optimization procedure involving the geometrical and structural definition of its main components, which are skins, stringers and frames, floor beams,

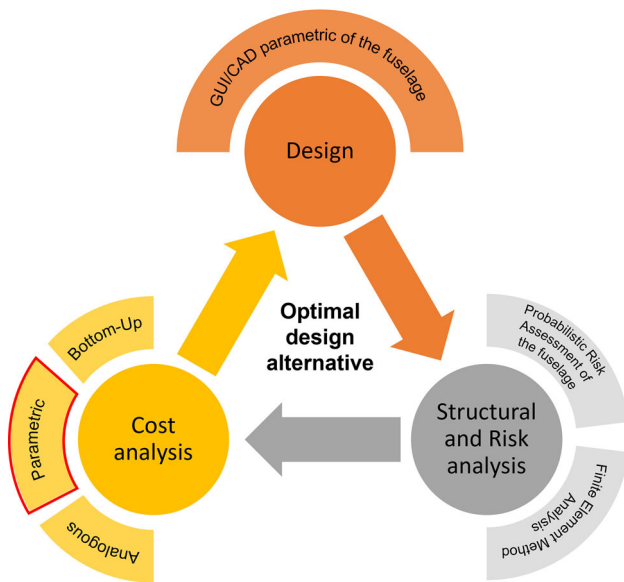


Fig. 1 SCERLOC design vision

pressure bulkhead, window frame and fittings. The MASCOT project belongs to the general framework named SHERLOC [20] (Structural Health monitoring, manufacturing and Repair technologies for Life management Of Composite fuselage) which design vision involves three main modules as can be seen from Fig. 1: (1) the design module that provides the GUI/CAD parametric representation of the fuselage; (2) the structural and risk analysis module that exploits two modules coming from SHERLOC i.e. the Finite Element Analysis (FEM) and Probabilistic Risk Assessment (PRA) to verify the structural requirements and manage the risks associated with the activities on the fuselage, respectively; finally (3) the cost analysis module that estimates the cost related with the manufacturing processes of the fuselage. It includes three sub-modules, i.e. Analogous, Parametric and Bottom-Up cost estimation methods.

Therefore, the SHERLOC design loop exploits data coming from all modules at each iteration of the design process to find the optimal design alternative that meets the requirements of all fields involved in the design process.

The presence of different cost estimation methods makes MASCOT more flexible enabling its use under different circumstances, i.e. the most appropriate cost estimation method should be chosen according to the requirements and data availability. It is worth noting that the three cost estimation methods require different inputs to estimate the cost of a new product achieving a different level of accuracy. In particular, the first two methods have a lower level of accuracy than the third ones because they use historical data to estimate the cost of a new product but require less information about product and manufacturing processes, therefore, are more appropriate to estimate the cost of product in the early phase of design.

While the last method, i.e. Bottom-Up, allows making a more accurate cost estimation but it requires detailed information about the product and its manufacturing process. Thus, for example in light of the above considerations, the analogous method could be used for a quick cost estimation when little data about the new product are known, and a high level of accuracy is not required. The parametric cost estimation could be used to analyse the effects of the variation of some parameters related to the product and process on its cost under specific assumptions, e.g., the new component must belong to the same family as the historical ones (for example they must share the same material, geometry and manufacturing process). Finally, the Bottom-Up approach could be used to get a detailed cost estimation under less stringent assumptions than the parametric method since the information required by the method allows to estimate the cost of the new product considering different materials, geometry and manufacturing processes.

For the specific purposes of this paper only the parametric cost estimation module developed for MASCOT project is described.

The remind of the paper is organised as follows: Sect. 2 reports a literature about cost estimation methods; Sect. 3 describes the proposed methodology in detail; Sect. 4 provides a numerical example to validate the proposed methodology, after which Sect. 4.1 shows how the reliability of the presented reduced regression model can be quantified, and finally, Sect. 5 presents the conclusions.

2 Literature review

As detailed in [21], three different cost estimation methods are usually adopted in the aeronautical field: analogue, parametric and bottom-up. Analogue models estimate cost by analogy with similar existing products, identifying the closest element among many available in historical data. An example of the application of the analogous approach to estimate the cost of space missions and maintenance for new aircraft is given in [22] and [23] respectively. Parametric models are also based on available data, but the cost estimation is governed by cost estimation relationships (CERs) between the cost and certain influential parameters [24]. One example of the parametric approach is given in [25], where a parametric equation describing the manufacturing cost of a metal plate structure, as a function of the plate geometry, was used to optimise the design of the plate to minimise manufacturing cost. The bottom-up approach estimates the total cost by identifying and dimensioning all components and detailing all activities involved. An example of a bottom-up approach applied to a composite aircraft part, with some simplified assumptions, is given in [16] and [26]. The analogous cost estimation method due to its characteristics is simple to use

and very fast but on the other hand it provides a rough cost estimate. Moreover, it is not possible to delve deeper into the relationships among all the features that affect the cost. In contrast, the bottom-up cost estimation method, due to its characteristics, enables the user to make a very detailed estimate, but on the other hand, it requires a significant amount of information and a deep knowledge of the manufacturing process. Moreover, it requires a great deal of effort in terms of time and money, and it is necessary to build a different model for each alternative design. Therefore, it is not suitable for use in the early phases of design when little information is known about a new product. Instead, the literature suggests the parametric approach as the best method to use in the early phase of the design phase [27–29] when little information about new product is known and, due to its characteristics, it can be used to perform sensitivity analyses, i.e., what happens to the cost if a particular feature is changed? Therefore, it is more straightforward to analyse several alternative designs. This is possible because the parametric method considers several cost drivers (main factors influencing the cost) in its formulation.

The parametric method can be traced back to the work of Wright when he first proposed the learning curve [30]. The technique is widespread within the aerospace industry, even if its application has often involved relatively simple statistical approaches. In [31] the manufacturing cost of commercial aircraft was carried out by a parametric approach coupled with neural networks and linear regression models created via Principal Component Analysis (PCA); the data of 17 aircraft projects were collected and seven global factors, such as total weight and maximum thrust, were identified as the main factors influencing the cost, i.e. cost drivers. In [32], cost drivers for material cost related to skin, sheets, extrusions, rivets, and fabrication cost related to skin, stringers, frames and cleats, and assembly cost involving manual/automatic riveting, were detailed in a parametric approach to estimate the cost of several aircraft panels; the cost estimation procedure was then coupled to a simplified structural analysis to obtain the optimal design. A simple but useful description of the generic parametric procedure, as well as its advantages and disadvantages, can be found in [33]. As detailed by [27], a difficult aspect of parametric cost estimation is the actual CER formulation itself; the cost model's robustness and reliability is heavily dependent on the type and size of the database used. Furthermore, the quality of the representative regression models also affects the reliability of the cost estimation tool. A further complicating factor that is related to composites, is that the manufacturing of composite components is more labour-intensive than traditional materials. One consequence of this, is that the manufacturing cost of a composite component is no longer simply related to the empty weight of the component, but needs to be estimated with less empirical relationships [34].

The parametric approach has some general features that make it applicable to a wide range of applications, especially for systems in the early phases of development. For instance in [35] some CERs were developed for estimating the manufacturing cost of commercial satellite launch vehicles. In [36] a parametric approach was adopted to evaluate the cost effectiveness of alternative infrastructure investments in civil engineering; the study was limited to roadway construction and the model was set in terms of a few data variables, no geometrical or physical properties were taken into account.

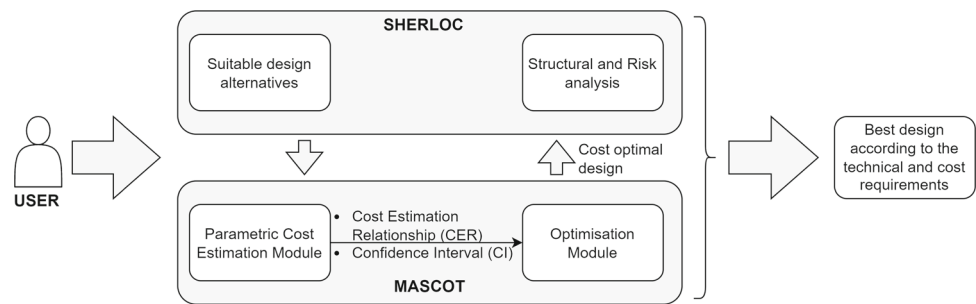
Although the approaches taken in the past by the research community to estimate production costs have often been very accurate, the research community has not yet considered the reliability of its cost estimates in the context of composite materials. There are many sources of uncertainty that heavily influence the reliability of an estimate for the manufacturing cost of a new composite aircraft component; databases for previous similar components might be incomplete, measurements of dimensions could be incorrect, prices of subcomponents could fluctuate for unknown reasons, the manufacturing processes used can change over time even for the same component etc. Therefore, it is important to consider these sources of uncertainty during the cost estimation process and quantify the reliability of the cost estimates obtained for new composite aircraft components. This would allow engineers to allocate costs in a more reliable manner, avoiding errors in budget definition and enabling resources to be allocated more effectively.

Thus, estimating the cost of a new product is a difficult process [37]. In the initial design phase, there are several hidden cost items that are not always easy to quantify, such as many of the details of the production process, e.g., the number and type of machines, indirect costs, labour, etc. Using the parametric method enables these critical issues to be overcome, since the total cost of the component is considered, which is obtained from historical data that indirectly considers all cost sources, allowing for a more complete view.

In this regard, this paper proposes a new cost estimation procedure that allows the full potential of the regression method to be exploited, presenting a procedure that identifies the most suitable parameters to be used. In addition, it allows the reliability associated with production cost estimation to be quantified, helping engineers to use this decision-making tool in the optimisation cycle that starts with a design, moves on to production cost estimation with a certain level of reliability, and finally modifies the design if the cost or reliability does not meet the objectives. This provides both a more accurate picture of how each independent variable affects cost, allowing for more accurate project cost optimisation, and the avoidance of errors in budget setting, allowing for more effective resource allocation.

The main novelties of this research are as follows (i) the development of a parametric modelling approach that enables

Fig. 2 Identification of the best design alternative—design iterative loop



the efficient choice of cost factors, considering the limitations of linear regression; and (ii) the analysis of the reliability of the cost estimate based on the confidence interval obtained from the regression model created.

3 Methodology

The proposed methodology provides a costing method to use in the early design phase of a new product to choose the best design alternative among the suitable ones according to the structural and cost requirements. Furthermore, a configuration of the product that minimizes the probability of going over-budget is provided by means of a reliability analysis. Nevertheless, this must be first confirmed by the structural and risk analysis before its final acceptance, as shown in the design iterative loop of Fig. 2.

The parametric cost estimation method was used to estimate the cost of a new component. This method estimated the cost as a function of one or more independent variables that are believed to be strongly correlated with it and, and so are named “cost drivers”. In this work, multiple linear regression was used to derive the cost estimation relationship (CER). This, in its general form, appears as in Eq. (1):

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ik} + \varepsilon_i \quad i = 1, \dots, n \tag{1}$$

where Y is the dependent variable (the variable whose value is to be estimated, i.e., the cost in this study), X is the independent variable (the variable used to estimate the cost of the dependent variable, i.e. “cost drivers”), n is the number of observed values, k being the number of independent variables/cost drivers, β are the regression coefficients and ε_i is a term accounting for random error. As can be noted from Eq. (1) a database of n historical data values is necessary to compute the β coefficients establishing in this way the relationship between the cost and cost drivers. Thus, the above equation can be written in matrix form as:

$$Y = X\beta + \varepsilon \tag{2}$$

where:

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1k} \\ x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix} \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \tag{3}$$

where the matrix X and vector Y compose the historical database. A linear regression model can be fitted to the data shown in Eq. (3):

$$\hat{y}_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} \dots + \beta_k x_{ik} \quad (i = 1, 2, \dots, n) \tag{4}$$

where \hat{y}_i is the predicted cost for the i 'th part x_i , and $\beta_0, \beta_1, \dots, \beta_k$ are regression coefficients.

The main steps of the cost estimation module are summarized in flowchart of Fig. 3, where can be seen that some novel features are used to fully exploit the performance of the linear regression: (i) the homoskedasticity assumption is tested and mitigated if necessary; (ii) only variables with a strong linear relationship and effect on the cost are used in the regression model by the use of the Pearson coefficient; (iii) stepwise forward procedure is used to identify the most useful cost drivers. Following a detailed description of the five steps shown in Fig. 3 is given.

Hypothesis The parametric cost estimation method relies on a CER to estimate the cost of a product. CER establishes a link between the cost (Y) and cost drivers (X_k) and, it is inferred from the historical values included in a database. Thus, the estimation of the cost of a new product by means of this method requires that the new product must be similar to the ones included in the database used to create CER, i.e. the new component must belong to the same family of the historical ones (for example they must share the same material, geometry and manufacturing process).

Step 1 In the first step, a database is created for the aircraft component under study, namely the X matrix and the Y vector of Eq. (3). To create the database, information is gathered on the cost of other similar components, as well as the values of the cost drivers (i.e., the characteristics that affect the cost

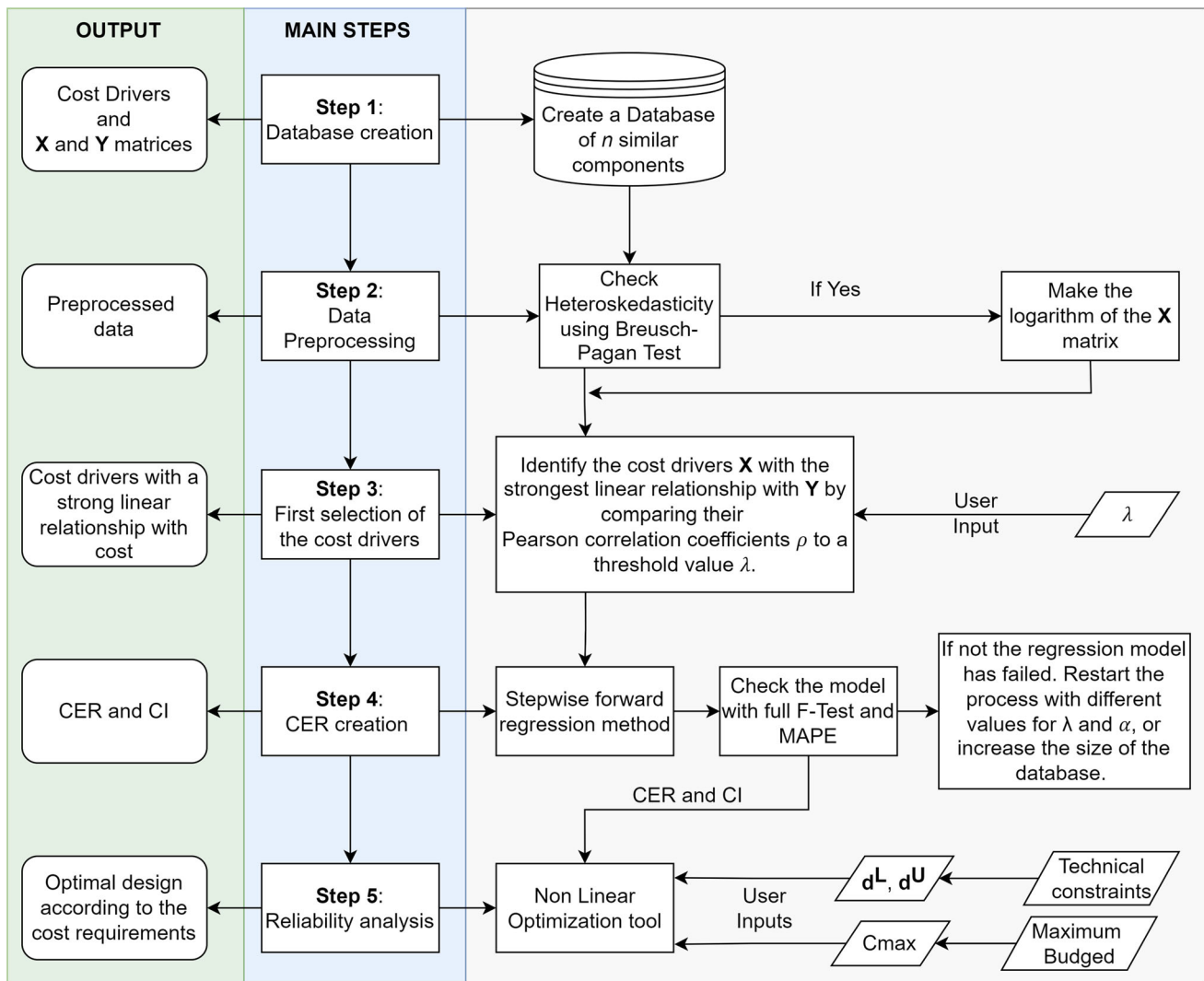


Fig. 3 Flowchart of the proposed methodology for manufacturing cost estimation

of producing the product, e.g., the production time needed, the curing time, etc.) for each of these similar components. Cost drivers included in X matrix could be production or design parameters. Typically, cost drivers are chosen based on expert experience and/or data reported in the literature. It is important that each identified cost driver is measurable and available at an early design phase. Regarding the vector Y , it includes the known costs of the components included in the database. It is worth noting that Y represents the cost of item e.g., it could be the total manufacturing cost, labour cost, testing cost, etc. Only complete observations must be included in the database, i.e. observations in which both the cost and all cost driver values are known. Then the relevant range for each cost driver is evaluated. This includes all values, from smallest to largest, for each cost driver used in the development of the regression model. It is worth noting that the product cost (Y) included in the database could

be its total production cost, thus, parametric cost estimation can consider indirectly all cost items related to the production of the product and overcome the uncertainty related to the knowledge of all cost items in the early design phase of the product. Two examples taken from literature are provided for a better comprehension of this step. *Example 1.* The authors in [31] estimated the general aviation development cost using collected data across 17 similar aircraft past projects with seven cost drivers that are: maximum take-off weight (tons), Mach number (Ma), maximum range (km), maximum thrust (kips), maximum ceiling (m), maximum oil load (L) and the length of the fuselage (ft). Thus, for each of the 17 past aircraft projects the vector Y includes the value of the general aviation development cost while the matrix X includes the values of the seven cost drivers. *Example 2.* The authors in [25] estimated the manufacturing cost of a plate structure using collected data across 100 similar historical

plates and, six cost drivers that are: outer width (m), outer length (m), outer radius (m), inner width (m), inner length (m) and thickness (m). Thus, for each of the 100 historical plates the vector \mathbf{Y} includes the value of the manufacturing cost while the matrix \mathbf{X} includes the value of the six cost drivers.

Step 2 Since linear regression assumes the absence of heteroskedasticity, the second step tests and mitigates its significance in the database. There are several methods to test whether the hypothesis of homoskedasticity is satisfied: the Breusch Pagan test, the White test, the Goldfeld-Quandt test, the Harvey Godfrey test and so on. Of these, the Breusch Pagan test was used in our case study due to its widespread use and its good compromise between simplicity and power [38]. If heteroskedasticity is found to be significant, its effect can be mitigated by taking the log of the matrix \mathbf{X} in Eq. (3), i.e., $\mathbf{X} = \ln(\mathbf{X})$, before creating the regression model seen in Eq. (4) [39] [40].

Step 3 The cost drivers (\mathbf{X}) with the strongest relationship with the cost are identified based on their Pearson correlations coefficients with respect to the cost (\mathbf{Y}). The Pearson correlation coefficient is a statistical tool able to provide an indication of how much two variables are linearly dependent on each other; such dependence does not imply the existence of a cause—effect relationship between them, but simply implies that the variability of one of them is related to the variability of the other. The Pearson correlation coefficient ($\rho(a, b)$) between two variables, a and b , can assume value in range $[-1; 1]$, where $\rho = \pm 1$ means a perfect positive (+ 1) or negative (− 1) linear correlation between a and b , while $\rho = 0$ means that no linear relationship exists between a and b [40].

Therefore, a user-defined threshold coefficient λ is used to determine which cost drivers exclude based on the strength of their linear relationship with cost. Indeed, if $\rho(\mathbf{X}_j, \mathbf{Y}) \geq \lambda$, where \mathbf{X}_j is the j 'th column of \mathbf{X} , then the j 'th cost driver is accepted for the next selection. It is worth to note that closer is λ to 0 less accurate will be the prediction because cost drivers with a weak linear relationship with the cost will be considered too.

Step 4 The stepwise forward method is used to identify the most important cost drivers (i.e., those linearly independent of each other) overcoming in this way the multicollinearity problem among the cost drivers included in the model. Indeed, the stepwise forward regression method enables the creation of an accurate regression model without examining all possible models [40]. The starting point of this method is a regression model that only includes one cost driver: usually the one with the highest Pearson coefficient with the cost. Following this, the cost driver with the next highest Pearson coefficient is added. Its effect on the regression model is determined via partial F test and adjusted R^2 ; if the test succeeds and the adjusted R^2 of the model is improved, the

newly added cost driver is judged to have improved the model and it is kept in the model; otherwise, the newly added cost driver is judged to have worsened the model and it is removed from the model. This procedure is repeated for all the cost drivers which have a Pearson coefficient with the cost higher than a user-defined threshold value, λ . Therefore, the final model will only include those p cost drivers that are deemed necessary to ensure the accuracy of the model; reducing the complexity of the model and ensuring that the problem of multicollinearity is mitigated. Three types of test are performed in this fourth step: (i) *F-Test* is performed on the incomplete regression model to judge if the addition of a variable has improved its performance; (ii) *Overall F-Test* is performed on the final regression model to judge if the regression relationship between the cost and the final set of cost drivers is statistically significant; (iii) *Mean Absolute Percentage Error (MAPE)* is performed on the final regression model to judge if the average of the absolute percentage errors of predictions where error is defined as actual or observed value minus the predicted value is acceptable. Finally, if the regression model is accepted, the $100 \times (1 - \alpha)$ percent confidence interval (CI) for a new cost prediction \hat{y}_{new} (also called the prediction interval) can be calculated via the following equation:

$$\hat{y}_{new} \pm t_{\alpha/2, n-(k+1)} \sqrt{MSE(1 + \mathbf{x}_{new}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_{new}')} \quad (5)$$

where \mathbf{X} is a historical database of n parts and p cost drivers, and \mathbf{Y} is a list of historical manufacturing costs corresponding to the n parts in \mathbf{X} , as seen in Eq. (3), MSE is the Mean Squared Error of the regression model, i.e.:

$$MSE = \frac{1}{n - (k + 1)} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (6)$$

where \hat{y}_i is the predicted cost from the regression model for the i 'th part in \mathbf{X} . \mathbf{x}_{new} in Eq. (5) is a vector containing the value of the cost drivers of a new part for which a cost \hat{y}_{new} needs to be predicted:

$$\mathbf{x}_{new} = [x_{new1} x_{new2} \cdots x_{newk}] \quad (7)$$

$t_{\alpha/2, n-(k+1)}$ in Eq. (5) can be found from the Student's t -distribution. For a confidence interval of 95%: $t_{\alpha/2, n-(k+1)} \approx 1.96$.

Brief descriptions of the three tests executed in this step are given in sections 3.1, 3.2 and 3.3.

3.1 F-test

The partial F test is a method used to evaluate the contribution provided by each cost driver to the model. It determines the

contribution to the regression sum of squares made by each cost driver after all the other cost drivers have been included in the model. Therefore, the new cost driver is included only if it significantly improves the model [40]. To verify the significance of the contribution given by the j 'th cost driver to the model, a statistical hypothesis test is executed considering the following hypotheses:

$H_{j,0}$: The j 'th cost driver does not significantly improve the model.

$H_{j,1}$: The j 'th cost driver does significantly improve the model.

The F statistic has a Fisher distribution with 1 and $n-k-1$ degrees of freedom. Given a user-defined significance level α , the decision rule is the following:

$$\text{reject } H_{j,0} \text{ if } F_j > F_{critical} \quad (8)$$

The value $F_{critical}$ is obtained by Fisher's distribution tables and is equal to: $F_{crit} = F^{-1}(1-\alpha; df_1, df_2)$, where F^{-1} is the inverse cumulative distribution function of the F-distribution, df_1 and df_2 are degrees of freedom, and α is a user-defined parameter.

3.2 Overall F-test

The overall F test is used to determine whether the regression relationship between the cost and the final set of cost drivers is statistically significant [40]. It is a statistical hypothesis test and for a regression model with p cost drivers the hypotheses are:

H_0 : There is no linear relation between the dependent and the independent variables.

H_1 : There is a linear relation between the dependent variable and at least one of the independent variables.

The F statistic has a Fisher distribution with p and $n-p-1$ degrees of freedom. Given a defined significance level α , the decision rule is the following:

$$\text{reject } H_0 \text{ if } F > F_{critical} \quad (9)$$

3.3 MAPE

A quality measure of a regression model is given by the MAPE. It is the mean or average of the absolute percentage errors of predictions where error is defined as actual or observed value minus the predicted value [38]. If n observations are given, the MAPE is computed by Eq. (10), in which

y_i and \hat{y}_i are the actual cost and the predicted cost of the i 'th observation respectively.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100 \quad (10)$$

Step 5 The parametric cost estimation method allows to create a CER capable to link the cost with the cost drivers, thus, once developed CER allows to make a "what if" analysis, i.e., it can assess how the cost of the product changes as the value of one or more cost drivers change [28]. In this work this ability is used to make a reliability analysis with the aim to find the product configuration that minimise the probability of going over-budget, i.e. named with \mathbf{d} the vector of design variables of product that can be modified ($\mathbf{d} \subseteq \mathbf{X}$) and with C_{max} the maximum budget, it is possible to set an optimisation problem that finds the value of \mathbf{d} that minimises the probability of going over-budget with a reliability of $R_{required}$:

$$\text{Minimise : } 1 - F(C_{max}, \text{Cost}_{\text{manufacturing}}(\mathbf{d}), CI(\mathbf{d})/2)$$

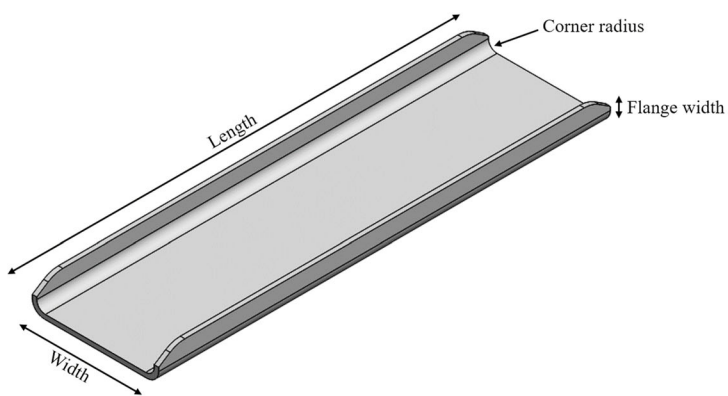
$$\text{Subject to : } 1 - F(C_{max}, \text{Cost}_{\text{manufacturing}}(\mathbf{d}), CI(\mathbf{d})/2) \leq 1 - R_{required}$$

$$\mathbf{d}^L \leq \mathbf{d} \leq \mathbf{d}^U$$

where $F = F(X, \mu, \sigma)$ is the Cumulative Distribution Function (CDF) of a normal distribution evaluated at X with a mean of μ and standard deviation σ ; $\text{Cost}_{\text{manufacturing}}$ is calculated using CER obtained in step 4; CI is the 95% confidence interval calculated using the procedure described in step 4. Instead, \mathbf{d}^L and \mathbf{d}^U are the technical constraints, i.e., are the lowest and highest value that the elements of \mathbf{d} can assume, respectively.

4 Numerical example: Boeing 787 stanchion

To demonstrate the proposed methodology, it was applied to a composite support reinforcement, also known as a stanchion, from the fuselage of a Boeing 787. The geometry of a stanchion is shown in Fig. 4a, while its location within the Boeing 787 fuselage is illustrated in Fig. 4b. The database used to demonstrate the effectiveness of the presented estimation method contains details of 176 similar stanchions used in the Boeing 787, due to the privacy issues it is not possible to report the real data; thus, they were altered by multiplying the real ones by a scale factor. Of these stanchions, length, corner radius, width, flange width and number of composite layers used in the strut were reported. Table 1 shows the max-



a)

b)

Fig. 4 The figure represents: **a** A stanchion from the fuselage of a Boeing 787 and **b** Location of the stanchion in the fuselage of a Boeing 787 highlighted by a dashed red circle [41]

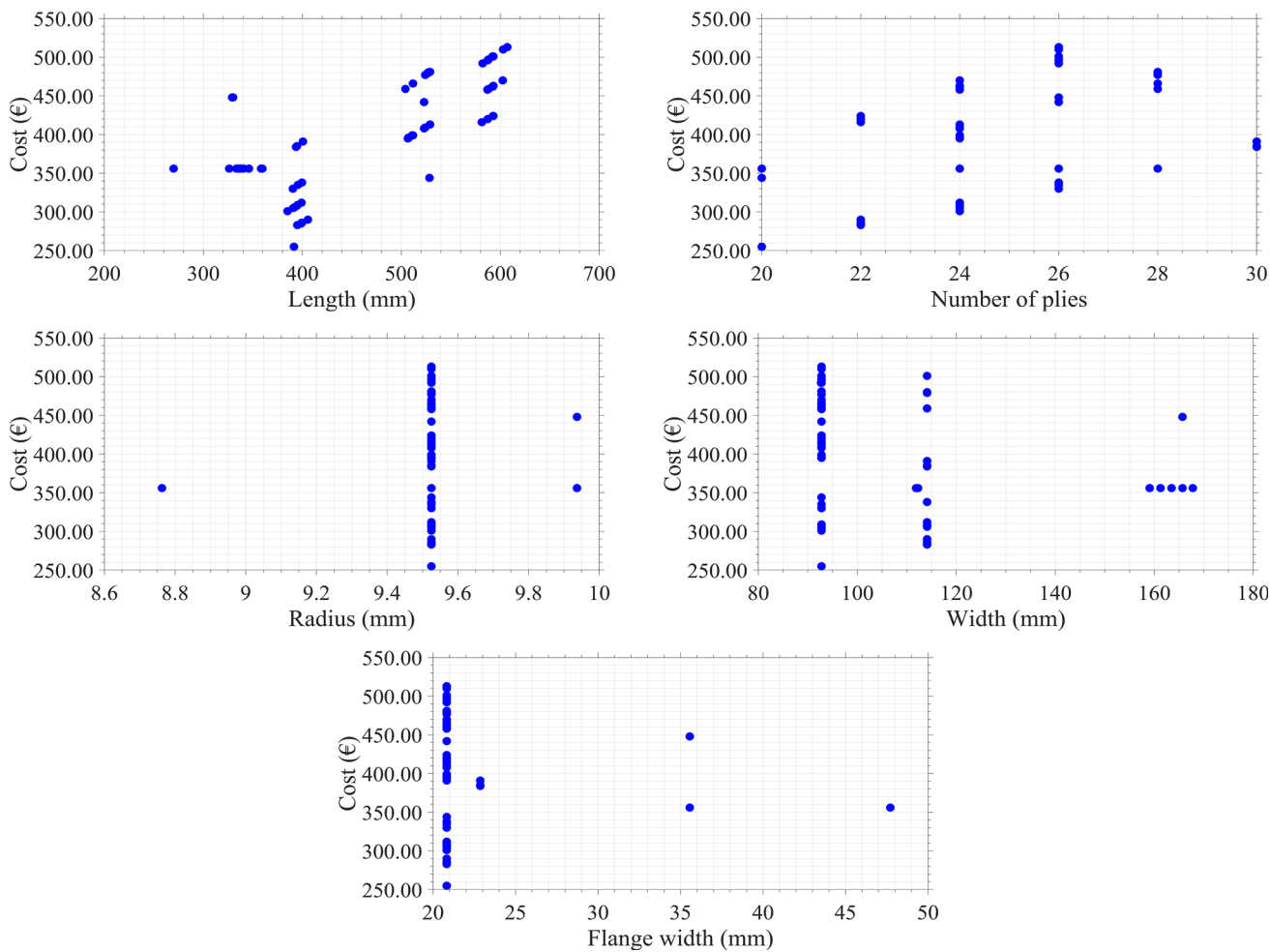


Fig. 5 The database for the stanchion

Table 1 Stanchion properties from the database. The minimum, mean, and maximum values are shown for each property

Beam property	Minimum	Mean	Maximum
Length (mm)	280	473	607
Corner radius (mm)	8.76	9.55	9.94
Width (mm)	92.7	109	168
Flange width (mm)	20.8	23.9	399
Number of composite plies	20	24.7	30
Manufacturing cost (€)	255.00	399.07	513.00

imum, minimum and main values of the individual variables.

Figure 5 shows the raw data from the database.

Based on this database, the multiple regression model for this numerical example has the following form:

$$\begin{aligned} \text{Cost}_{\text{manufacturing}} = & C_0 + C_{\text{Length}} \text{Length} + C_{N\text{-plies}} N.\text{plies} \\ & + C_{\text{Radius}} \text{Radius} + C_{\text{Width}} \text{Width} \\ & + C_{\text{Flange_width}} \text{Flange width} \end{aligned} \quad (11)$$

The procedure outlined in Sect. 3 is used to create a regression model of the form seen in Eq. (11) using the stanchion database. During preliminary tests, the *p*-value from the Breusch-Pagan test for heteroskedasticity (Sect. 3, step 2) was found to be 4.76×10^{-6} . This value is less than the threshold of 0.05, thus indicating that heteroskedasticity is significant in the database. Therefore, the *x*-data seen in Fig. 5 is logged and the multiple regression model for this numerical example will now have this form:

$$\begin{aligned} \text{Cost}_{\text{manufacturing}} = & C_0^* + C_{\text{Length}}^* \ln(\text{Length}) \\ & + C_{N\text{-plies}}^* \ln(N.\text{plies}) \\ & + C_{\text{Radius}}^* \ln(\text{Radius}) \\ & + C_{\text{Width}}^* \ln(\text{Width}) \\ & + C_{\text{Flange_width}}^* \ln(\text{Flange width}) \end{aligned} \quad (12)$$

where the superscript * in the regression coefficients indicates that the regression coefficients in Eq. (12) are different than the regression coefficients in Eq. (11).

Following the procedure, the Pearson correlations coefficients were computed, and they are provided in Table 2. It can be seen that the cost is heavily correlated with Length, Number of plies (N. plies), and Width. The cost is weakly correlated with the Radius, and the Flange width. Several of the independent variables are also correlated with each other, such as Length and Width, Length and Flange width.

According to the procedure a threshold value for Pearson coefficient was arbitrarily set to 0.3 ($\lambda = 0.3$). Thus, only cost drivers which absolute value of the Person coefficient

was greater than 0.3 were selected to the next step of the procedure; thus, the following ones were selected: Length, N. plies and Width.

As part of the proposed procedure, stepwise regression was used to determine the independent variables that have the greatest influence on the cost. Thus, several regression models were created adding one cost driver per time and evaluating its usefulness according to the R^2 value and partial F-Test, as shown in Table 3. The value F_{critical} was obtained by Fisher's distribution tables considering $\alpha = 0.1$, $n = 176$, while *k* is equal to 2 and 3, for the 2nd and 3rd regression model, respectively.

As can be seen from Table 3 all cost drivers identified in the previous step (Length, N. plies and Width) were included in the final model as each of them significantly improve its performance, indeed each of them increased the value of R^2 and passed the partial F-test.

The overall F-test was executed on the 3rd regression model to judge if the regression relationship between the dependent variable and the final set of independent variables is statistically significant. To this purpose the F statistic was computed and compared with the critical value of F obtained by Fisher's distribution tables considering $\alpha = 0.1$, $n = 176$ and $k = 3$. The F_{critical} was found equal to 2.705 and as it was lower that the value of $F = 201.8$ the test was considered passed. Thus, the 3rd regression model was assumed as a final reduced regression model which contains fewer independent variables than that seen in Eq. (12).

The statistics of the reduced regression model can be seen in Table 4, while the statistics of its regression coefficients can be seen in Table 5. The R^2 and R_{adj}^2 for the reduced model are 0.779 and 0.774, respectively. Moreover, the MAPE was computed for the reduced model, and it is equal to 6.56%, i.e., in average the prediction error of the model is about 6.5%.

4.1 Reliability of the cost estimate

In this section, the reliability of the reduced regression model presented in the previous section is quantified. The reduced regression model, with the values for the regression coefficients seen in Table 5, is:

$$\begin{aligned} \text{Cost}_{\text{manufacturing}} = & -2825.8 + 314.25 \ln(\text{Length}) \\ & + 249.81 \ln(N.\text{plies}) \\ & + 106.22 \ln(\text{Width}) \end{aligned} \quad (13)$$

This reduced regression model is plotted in Fig. 6 for each of the individual variables in the database. 95% confidence intervals were calculated for this reduced regression model using the procedure described in Sect. 3 at step 4 and are also included in Fig. 6. The average 95% confidence interval over the range of *x*-values seen in Fig. 6 was ± 73.78 €, with

Table 2 Pearson correlation coefficients for the stanchion database

	Length	N. plies	Radius	Width	Flange width	Cost
Length	1.00	–	–	–	–	–
N. plies	0.06	1.00	–	–	–	–
Radius	– 0.16	0.47	1.00	–	–	–
Width	– 0.74	0.21	0.57	1.00	–	–
Flange width	– 0.68	– 0.03	0.13	0.80	1.00	–
Cost	0.78	0.46	– 0.04	– 0.37	– 0.26	1.00

Table 3 Stepwise forward procedure—intermediate regression models

	CER	R ²	F	F _{critical}
1st regression model	$Cost_{\text{manufacturing}} = C_0^{1st} + C_{\text{Length}}^{1st} \ln(\text{Length})$	0.574	–	–
2nd regression model	$Cost_{\text{manufacturing}} = C_0^{2nd} + C_{\text{Length}}^{2nd} \ln(\text{Length}) + C_{\text{N.Plies}}^{2nd} \ln(\text{N.plies})$	0.750	122.129	2.705
3rd regression model	$Cost_{\text{manufacturing}} = C_0^{3rd} + C_{\text{Length}}^{3rd} \ln(\text{Length}) + C_{\text{N.Plies}}^{3rd} \ln(\text{N.plies}) + C_{\text{Width}}^{3rd} \ln(\text{Width})$	0.779	21.969	2.705

Table 4 Statistics of the reduced model

Statistic	Value
R ²	0.779
R ² _{adj}	0.775
F-value	201.8
p-value	4.19 × 10 ^{−56}
MAPE	6.56%

Table 5 Coefficient statistics of the reduced model

Coefficient	Regression coefficient	p-value	Standard error
C [*] ₀	– 2825.8	2.08 × 10 ^{−30}	200.9
C [*] _{Length}	314.25	1.86 × 10 ^{−34}	20.28
C [*] _{N.Plies}	249.81	1.45 × 10 ^{−15}	28.40
C [*] _{Radius}	–	–	–
C [*] _{Width}	106.22	5.61 × 10 ^{−6}	22.66
C [*] _{Flange_width}	–	–	–

a maximum of ± 84.78 €, and a minimum of ± 71.57 €, representing significant levels of uncertainty.

Using Eq. (13) and the procedure described in Sect. 3, the manufacturing cost and 95% confidence interval for a new stanchion design with: Length = 525.35 mm, N. plies

= 28, radius = 9.5 mm, width = 140 mm, and flange width = 30 mm, are estimated to be 500.00 € and 67.25 € respectively. As shown in Fig. 7, this is equivalent to a normal distribution with a mean of 500.00 €, and a standard deviation of 33.63 €:

Given the significant level of uncertainty associated with the cost shown in Fig. 7, an engineer using the regression model in Eq. (13) cannot be certain that the actual cost of the new stanchion will be 500.00 €. If the engineer is assigned a budget of 500.00 € for the new stanchion (C_{max} = €500), then a 50% chance of going over this budget would be undesirable. To minimise the probability of going over-budget, and therefore enable engineers to allocate their resources in a more effective manner, the regression model seen in Eq. (13) can be used to optimize the cost of a new stanchion. To this purpose the optimisation problem defined in Sect. 3 at step 5 was used:

$$\text{Minimise : } 1 - F(500, \text{Cost}_{\text{manufacturing}}(\mathbf{d}), CI(\mathbf{d})/2)$$

$$\text{Subject to : } 1 - F(500, \text{Cost}_{\text{manufacturing}}(\mathbf{d}), CI(\mathbf{d})/2) \leq 0.01$$

$$\mathbf{d}^L \leq \mathbf{d} \leq \mathbf{d}^U \tag{14}$$

where $F = F(X, \mu, \sigma) = F(500, \text{Cost}_{\text{manufacturing}}(\mathbf{d}), CI(\mathbf{d})/2)$ is CDF of a normal distribution evaluated at X with a mean of μ and standard deviation σ . \mathbf{d} is a vector of three design variables (length, N. plies, and width) Cost_{manufacturing} is calculated using Eq. (13). CI is the 95% confidence interval calculated using the procedure described in Sect. 3. This optimisation procedure is designed to find the optimal combination of values for the three design variables such that the

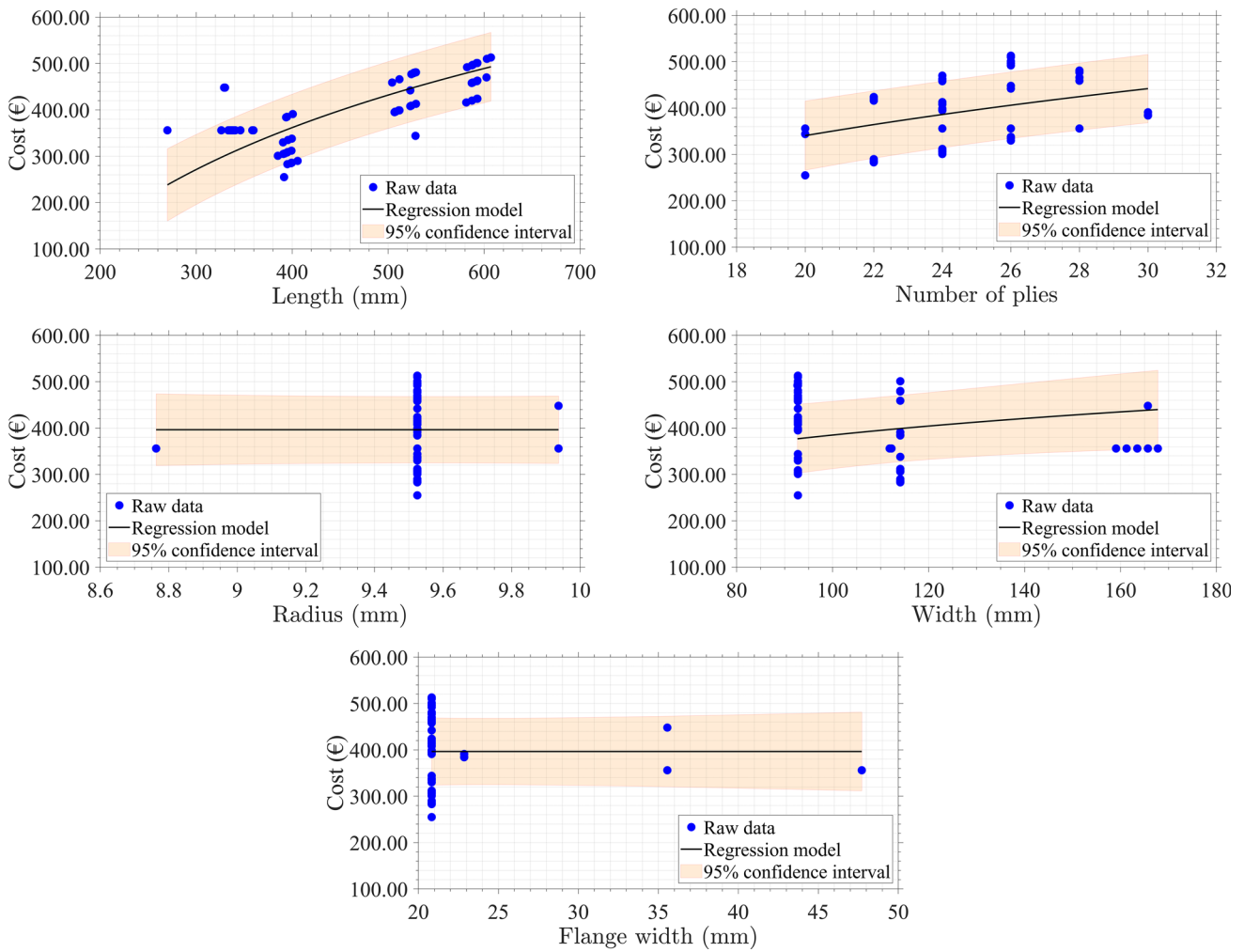


Fig. 6 Confidence intervals for the reduced regression model

probability of going over-budget is only 1% i.e., the reliability is 99% ($R_{required} = 0.99$).

The optimisation problem described in Eq. (14) is solved using Matlab® with the optimisation solver ‘fmincon’. The results can be seen in Table 6. The objective function F in Eq. (14) converges to 0.99 after four iterations. It is found that, to keep the cost of the design under or equal to 500.00 € with 99% reliability, the optimal design should have a length of 435.64 mm, the number of plies should be 28, and the width should be 116.29 mm.

This result represents an optimum from an economic point of view, as shows in the loop of Fig. 2 before its final acceptance, a new structural analysis must be carried out to check if the suggested dimensions satisfy the structural constraints too. The results in Table 6 underline that the proposed reliability-based approach can help engineers stay within budget with a high level of reliability; enabling them to allocate resources more effectively. In fact, if this reliability-based approach is not used, there is a 50% chance that the

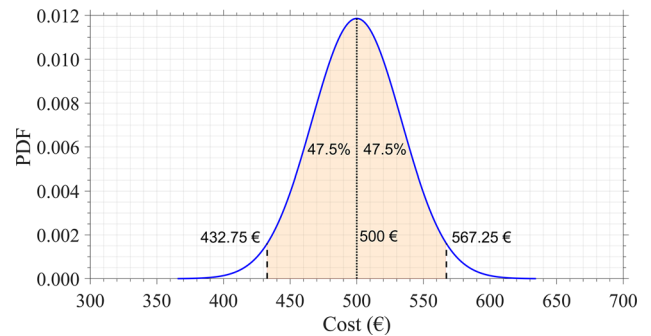


Fig. 7 Equivalent normal distribution for a new stanchion with an estimated cost of 500.00 € and a 95% confidence interval of 67.25 €

design will be over-budget, i.e., this would mean to look at the first row of Table 6. On the other hand, by using the approach presented in this paper, the chance that the design will be over-budget is significantly reduced to only 1%, i.e., the fourth row of Table 6. It is worth noting that the variables values shown in Table 6 are referred to a starting database

Table 6 Results of the optimisation solver ‘fmincon’ with the optimisation problem shown in Eq. (14)

Iteration	Length (mm)	N. plies	Width (mm)	Cost _{manufacturing} (€)	CI (€)	<i>F</i>
0	525.35	28	140.00	500.00	67.25	0.500
1	521.22	28	134.79	493.49	67.24	0.577
2	503.20	28	112.06	462.81	67.38	0.865
3	463.83	28	112.03	437.19	67.49	0.969
4	435.64	28	116.29	421.45	67.53	0.990

that was opportunely scaled for privacy issue, so the mechanical implications of these values are strongly limited by this facet.

5 Conclusions

This paper presents a novel methodology for efficiently choosing the cost drivers for composite aircraft components and to quantify the reliability associated with cost estimates. The estimation and the reliability of a production cost estimate is a well-known research topic especially at the early design phase of a new product when an accurate cost analysis could allow for improved cost management and help avoid unexpected costs during production. Indeed, a reliable estimate of costs is one of the key principles to maximise the company’s efficiency and, consequently, its profitability. Cost can be a deciding element in the design loop to identify the optimal product configuration according to technical and economic constraints. Obtaining a reliable cost estimate in the early design phase may be a heavy task due to the frequent scarcity of necessary information. This scarcity is even more significant when new materials are used in the production process, since it is the source of many of the uncertainties influencing the reliability of a cost estimate. The parametric approach developed in this work can overcome this issue by providing a cost estimate based on historical data and using a cost estimation relationship that links cost drivers (inputs) to the cost (output). The proposed methodology includes five steps, three of which focus on determining the best choice of cost drivers to fully exploit the ability of the linear regression while the last step involves a reliability analysis to identify a possible configuration of the product according to technical and economic constraints. The results show that the proposed reliability-based approach can lead to budget compliance with a high level of reliability, reducing the possibility of a budget overrun to only 1%. The results show a clear improvement over classical approaches that did not take reliability into account, allowing for a more efficient allocation of resources within a project. This research therefore offers an interesting starting point for the reliable estimation of new composite components in the aircraft industry. It

is worth noting that the structural reliability of the component was not considered in the cost model, e.g. the impact of different design alternatives against the failure rate of the component was neglected. However, even if this shortcoming represents a specific limitation for the presented study, it can be overcome by contextualising the methodology in the wider framework of the SHERLOC project, i.e. considering the structural and risk analysis module as shown in Figs. 1 and 2

Acknowledgements This work was supported by the European Commission under H2020 Framework Clean Sky JTI—project MASCOT, Grant Agreement n° 864154.

Funding Open access funding provided by Università degli Studi della Campania Luigi Vanvitelli within the CRUI-CARE Agreement.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Zhao, X., Verhagen, W.J.C., Curran, R.: Estimation of aircraft component production cost using knowledge based engineering techniques. *Adv. Eng. Inform.* **29**, 616–632 (2015)
2. Roy, R.: Cost engineering: why, what and how? Decision Engineering Report Series, no. Issue: 1, Cranfield University, 2003.
3. Moeeni, H., Javadi, M., Raissi, S.: Design for manufacturing (DFM): a sustainable approach to drive the design process from suitability to low cost. *Int. J. Interact. Des. Manuf.* **16**, 1079–1088 (2022)
4. Shehab, E., Ma, W., Wasim, A.: Manufacturing cost modelling for aerospace composite applications. *Adv. Concurr. Eng.* (2013). https://doi.org/10.1007/978-1-4471-4426-7_37/COVER
5. Adjoul, O., Benfriha, K., Aoussat, A.: Algorithmic strategy for optimizing product design considering the production costs. *Int. J. Interact. Des. Manuf.* **13**, 1313–1329 (2019)

6. Etienne, A., et al.: Cost engineering for variation management during the product and process development. *Int. J. Interact. Des. Manuf.* **11**, 289–300 (2017)
7. Luiza, D., Daniela, G., Alexandru, E.: Modeling and estimation of the product time and cost in manufacturing system-market relationship. *Int. J. Interact. Des. Manuf.* **8**, 277–282 (2014)
8. Chen, S., Li, D., Xiang, J.: A novel composite design optimisation method for minimized manufacturing cost with improved performances. *Appl. Compos. Mater.* **29**, 1479–1505 (2022)
9. Degenhardt, R., et al.: Future structural stability design for composite space and airframe structures. *Thin-Walled Struct.* **81**, 29–38 (2014)
10. Feng, Y., He, Y., Tan, X., An, T., Zheng, J.: Investigation on impact damage evolution under fatigue load and shear-after-impact-fatigue (SAIF) behaviors of stiffened composite panels. *Int. J. Fatigue* **100**, 308–321 (2017)
11. Bisagni, C., Dávila, C.G.: Experimental investigation of the post-buckling response and collapse of a single-stringer specimen. *Compos. Struct.* **108**, 493–503 (2014)
12. Fera, M., et al.: Economic and environmental sustainability for aircrafts service life. *Sustainability* **12**, 10120 (2020)
13. Pinto, J.P.C.: Economic comparison between materials in the aerospace industry: University of Lisbon (2017).
14. Shama Rao, N., Simha, T.G.A., Rao, K.P., Ravi Kumar, G.V.V.: Carbon composites are becoming competitive and cost effective. White Paper (2018)
15. Choi, J.W.: Architecture of a knowledge based engineering system for weight and cost estimation for a composite airplane structures. *Expert Syst. Appl.* **36**, 10828–10836 (2009)
16. Hagnell, M.K., Åkermo, M.: A composite cost model for the aeronautical industry: methodology and case study. *Compos. B Eng.* **79**, 254–261 (2015)
17. Schubel, P.J.: Cost modelling in polymer composite applications: case study—analysis of existing and automated manufacturing processes for a large wind turbine blade. *Compos. B Eng.* **43**, 953–960 (2012)
18. Sun, G., Huo, X., Wang, H., Hazell, P.J., Li, Q.: On the structural parameters of honeycomb-core sandwich panels against low-velocity impact. *Compos. B Eng.* **216**, 108881 (2021)
19. MASCOT (Modular multilevel cost Analysis Software for COMposiTe smart fuselage). <https://www.mascotproject.eu/>.
20. SHERLOC (Structural HEalth monitoring, manufacturing and Repair technologies for Life management Of Composite Fuselage). <https://sherloc-project.com/>.
21. Curran, R., Raghunathan, S., Price, M.: Review of aerospace engineering cost modelling: the genetic causal approach. *Prog. Aerosp. Sci.* **40**, 487–534 (2004)
22. Shishko, R.: Developing analogy cost estimates for space missions: A Collection of Technical Papers—AIAA Space 2004 Conference and Exposition **2**, 1530–1537 (2004).
23. Yang, P., Wang, L., Sun, F.: Maintenance cost prediction and evaluation of new airplane for general aircraft. *IET Conference Publications* **2018**, 1472–1477 (2018)
24. Hueber, C., Horejsi, K., Schledjewski, R.: Review of cost estimation: methods and models for aerospace composite manufacturing (2016). <https://doi.org/10.1080/20550340.2016.11546422,1-13>
25. Morse, L., Aliabadi, M.H., Mallardo, V.: Manufacturing cost and reliability-based shape optimisation of aircraft structures. *Int. J. Numer. Meth. Eng.* **123**, 2189–2213 (2022)
26. Morse, L., Cartabia, L. & Mallardo, V.: Reliability-based bottom-up manufacturing cost optimisation for composite aircraft structures. *Structural and Multidisciplinary Optimisation* **65** (2022).
27. Trivailo, O., Sippel, M., Şekerçioğlu, Y.A.: Review of hardware cost estimation methods, models and tools applied to early phases of space mission planning. *Prog. Aerosp. Sci.* **53**, 1–17 (2012)
28. NASA: NASA Cost Estimating Handbook. National Aeronautics and Space Administration, Washington, DC (2008).
29. Lillie, C.F., Thompson, B.E.: Parametric cost estimation for space science missions (2008). <https://doi.org/10.1117/12.7896157018,786-797>
30. Wright, T.P.: Factors affecting the cost of airplanes. **3**, 122–128 (2012). <https://doi.org/10.2514/8.155>
31. Chen, X., Huang, J., Yi, M.: Development cost prediction of general aviation aircraft projects with parametric modeling. *Chin. J. Aeronaut.* **32**, 1465–1471 (2019)
32. Castagne, S., et al.: A generic tool for cost estimating in aircraft design. *Res. Eng. Des.* **18**, 149–162 (2008)
33. Duverlie, P., Castelain, J.M.: Cost estimation during design step: parametric method versus case based reasoning method. *Int. J. Adv. Manuf. Technol.* **15**, 895–906 (1999)
34. Ali, R., Al-Shamma, O.: A comparative study of cost estimation models used for preliminary aircraft design. *Global J. Res. Eng.* **14**, 9–18 (2014)
35. Drenthe, N.T., Zandbergen, B.T.C., Curran, R., van Pelt, M.O.: Cost estimating of commercial smallsat launch vehicles. *Acta Astronaut* **155**, 160–169 (2019)
36. Swei, O., Gregory, J., Kirchain, R.: Construction cost estimation: a parametric approach for better estimates of expected cost and variation. *Transp. Res. B Methodol.* **101**, 295–305 (2017)
37. Caputo, F., Caterino, M., De Luca, A., Fera, M., Greco, A., Lamanna, G., Macchiaroli, R., Manco, P., Manzo, M., Perfetto, D.: Product and process integrated design to enhance smart manufacturing systems. *Commun. ECMS* **33**, 213–218 (2019)
38. Swamidass, P.M.: *Encyclopedia of Production and Manufacturing Management*. Springer, Berlin (2000)
39. Neter, J., Kutner, M.H., Nachtsheim, C.J., Wasserman, W.: *Applied linear statistical models* (1996).
40. Berenson, M.L., Levine, D.M.: *Basic business statistics: concepts and application*, 12 ed. Pearson (2011).
41. Ostrower, J.: Boeing nears end of further 787 revision. *FlightGlobal*. <https://www.flightglobal.com/boeing-nears-end-of-further-787-revision/88614.article> (2009).

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.