International Journal of Industrial Engineering Computations 3 (2012) 817–828

Contents lists available at GrowingScience

International Journal of Industrial Engineering Computations

homepage: www.GrowingScience.com/ijiec

A new approach for product cost estimation using data envelopment analysis

Fantahun M. Defersha[a*], **Adil Salam**[b] and **Nadia Bhuiyan**[b]

[a]*School of Engineering, University of Guelph, 50 Stone Road East, Guelph, Ontario, Canada, N1G 2W1*
[b]*Department of Mechanical and Industrial Engineering, Concordia University 1455 de, Maisonneuve W., Montreal, Quebec, Canada, H3G 1M8*

A R T I C L E I N F O

Article history:
Received 10 April 2012
Received in revised format 12 June 2012
Accepted June 22 2012
Available online 23 June 2012

Keywords:
Estimation
Parametric Methods
Non-Parametric Methods
Data Envelopment Analysis

A B S T R A C T

Cost estimation of new products has always been difficult as only few design, manufacturing and operational features will be known. In these situations, parametric or non-parametric methods are commonly used to estimate the cost of a product given the corresponding cost drivers. The parametric models use priori determined cost function where the parameters of the function are evaluated from historical data. Non-parametric methods, on the other hand, attempt to fit curves to the historic data without predetermined function. In both methods, it is assumed that the historic data used in the analysis is a true representation of the relation between the cost drivers and the corresponding costs. However, because of efficiency variations of the manufacturers and suppliers, changes in supplier selections, market fluctuations, and several other reasons, certain costs in the historic data may be too high whereas other costs may represent better deals for their corresponding cost drivers. Thus, it may be important to rank the historic data and identify benchmarks and estimate the target costs of the product based on these benchmarks. In this paper, a novel adaptation of cost drivers and cost data is introduced in order to use data envelopment analysis for the purpose of ranking cost data and identify benchmarks, and then estimate the target costs of a new product based on these benchmarks. An illustrative case study has been presented for the cost estimation of landing gears of an aircraft manufactured by an aerospace company located in Montreal, CANADA.

1. Introduction

Product cost assessment has become much more relevant than ever before because of the immense competition facing manufacturers. According to Layer et al. (2002), product cost assessment may be classified as pre-assessment, intermediate assessment and post-assessment. Pre-assessment is the process of estimating product costs using historical cost data of related products; intermediate assessments are required for controlling costs during the product development cycle. Cost accounting methods utilized to determine the actual cost incurred after the product is manufactured are classified as post-assessment. The final costs obtained by post-assessment will be used for future pre-assessment of related products. In make-to-order or engineer-to-order project oriented companies (e.g. aerospace), a large part of the product cost is defined during the pre-assessment phase as this cost information is essential for bidding purposes and to define selling prices and analyze future profitability. However, at

* Corresponding author. Tel: 1-(519) 824 4120 Ext. 56512
E-mail: fdefersh@uoguelph.ca (F. M. Defersha)

doi: 10.5267/j.ijiec.2012.05.009

this early design stage, only few and generic product characteristic are usually known. In such situations, parametric cost estimation methods have been very attractive as these methods seek to evaluate the cost of a product from parameters characterizing the product but without describing it in detail. Parametric methods use the relationship between the physical characteristics of the part, such as mass or volume, and the cost, but with little or no physical relationship to the process. In these methods, statistical criteria are utilized to identify the causal links and correlate costs and product characteristics in order to obtain parametric function with one or more variables (Foussier, 2006). For example, regression analysis have been widely utilized (Phaobunjong & Popescu, 2003; Dean, 2005). However, before the parameters are determined, parametric methods require predetermined parametric function to be chosen by the cost estimator that he/she believes will lead to good fit to the historical data. The choice of these functions also depends on the relative simplicity of the analysis required to determine the parameters. If for example the cost estimator chooses very complicated non-linear function with several unknown parameters, the determination of these parameters may require the use of complicated heuristic algorithms (see for example Zheng & Zhang, 2006; Chan et al., 2010).

Non-parametric methods, on the other hand, attempt to fit curves to the historic data without predetermined function. Among such methods is artificial neural network (ANN). ANNs have the ability to classify and extrapolate collections of data (Pandya & Macy,1996). ANN models accept shape-describing and semantic product characteristics as inputs and give as output the product cost (Bode, 1998, 2000). Zhang et al. (1996) and Zhang and Fuh (1998) demonstrated the use of ANN in estimating packaging costs based on product dimensions. Seo et al. (2002) apply ANN methods in life cycle costing during the conceptual design stage. A case study of manufacturing cost estimation of machined components in an automotive industry using ANN is presented in Cavalieri et al. (2004). Several papers report comparative studies between ANN and parametric regression methods (Cavalieri et al., 2004; Caputo & Pelagagge, 2008; Verlinden et al., 2008; Dura et al., 2009). These articles assert that ANNs show better cost estimation performance than regression analysis. In both parametric and non-parametric methods discussed above, it is assumed that the historic data used in the analysis is a true representation of the relationship between the product characteristics (cost drivers) and the corresponding costs. This is evidenced by the fact that the objective function used in the search of parameters or connection weight is the goodness of the fit of the input/output from these methods to that of the historic data. However, because of efficiency variations of the manufacturers and suppliers, changes in supplier selections, market fluctuations, and several other reasons, certain costs in the historic data may be too high whereas other costs may represent better deals for their corresponding cost drivers. Thus, it may be important to rank the historic data and identify the best, average, or worst benchmarks and provide different target costs of the product based on these benchmarks. The manufacturer then may strive to be able to produce the product at cost as close as possible to the cost estimate obtained using the best benchmark. When such cost appears not achievable, the manufacturer may determine its cost based on different benchmark such as the historically observed average performance level or higher.

In this paper, novel adaptation of cost driver and cost data is introduced in order to use data envelopment analysis (DEA) for the purpose of ranking cost data and identifying benchmarks, and then estimating the target costs of new product based on these benchmarks. The remainder of this paper is organized as follows. In Section 2, we provide an introduction to the early DEA models which are also used as the basis for the work presented in this paper. The process of adaptation of cost drivers and cost data in order to use DEA as a cost estimation tool is presented in Section 3 followed by implementation procedures in Section 4. An illustrative case study is provided in Section 5. In Section 6, we illustrate an empirical relationship that may exist between the DEA based approach proposed in this paper and the parametric or non-parametric methods. Discussion and conclusions are in Section 7.

2. Basic DEA models

As the main objective of this paper is to demonstrate new use of DEA as cost estimation tool, we did not attempt to provide detailed literature review of DEA in the way they are commonly used in evaluating performances of decision making units (DMUs). In this section, we simply provide the basic DEA models that we use in the proposed cost estimation method. A comprehensive review of DEA models can be found in Adler et al. (2002). DEA is originally proposed by Charnes et al. (1978) in order to measure the relative efficiency of homogeneous DMUs with multiple inputs and outputs. It is applied when there is no obvious unit price information for some or all of the inputs and the outputs to aggregate them into single equivalent input and single equivalent output, respectively. In DEA, relative efficiency of DMU is defined as the ratio of total weighted output to total weighted input. By considering a set of N homogeneous decision making units (DMU_n for $n = 1, 2, \cdots, N$) each having I number of inputs ($x_{i,n}$, for $i = 1, \cdots, I$) and O number of outputs ($y_{o,n}$ for $o = 1, \cdots, O$), the efficiency measure E_k for DMU_k isgiven by Eq. 1, where the weights $u_{i,k}$ and $v_{o,k}$ are non-negative and unknown until they are determined by the DEA procedure.

$$E_k = \frac{\sum_{o=1}^{O} v_{o,k} \cdot y_{o,k}}{\sum_{i=1}^{I} u_{i,k} \cdot x_{i,k}} \tag{1}$$

The weights $u_{i,k}$ and $v_{o,k}$ corresponding to DMU_k are determined in such way that the efficiency E_k of this decision making unit can be maximized subject to the following constraints. The first one is, when these weights are applied to all DMUs they should not provide any DMU with efficiency greater than one. The second one requires the weights to be non-negative. This problem can be formulated for DMU_k as a fractional linear programming mathematical model as follows.

Maximize (for a given k):

$$E_k = \frac{\sum_{o=1}^{O} v_{o,k} \cdot y_{o,k}}{\sum_{i=1}^{I} u_{i,k} \cdot x_{i,k}} \tag{2}$$

subject to

$$\frac{\sum_{o=1}^{O} v_{o,k} \cdot y_{o,n}}{\sum_{i=1}^{I} u_{i,k} \cdot x_{i,n}} \leq 1;\ \forall n \tag{3}$$

$$u_{i,k} \geq 0;\ v_{o,k} \geq 0;\ \ \forall i, \forall o \tag{4}$$

DMU_k will choose weights $u_{i,k}$ and $v_{o,k}$ so as to maximize its efficiency, given the constraints in Eqs. (2) and (3). The fractional linear programming described above can be translated into simple linear programming by multiplying both the numerators and denominators of the fractions with positive constant c and choosing the constant such that, $c \times \sum_{i=1}^{I} u_{i,k} \cdot x_{i,k} = 1$. The products of the constant c and the variables $u_{i,k}$ and $v_{o,k}$ can be replaced by new variables $p_{i,k}$ and $q_{o,k}$; respectively. The resulting linear programming is shown below.

Maximize (for a given k):

$$E_k = \sum_{o=1}^{O} q_{o,k} \cdot y_{o,k} \tag{5}$$

subject to

$$\sum_{o=1}^{O} q_{o,k} \cdot y_{o,n} - \sum_{i=1}^{I} p_{i,k} \cdot x_{i,n} \leq 0;\ \forall n \tag{6}$$

$$\sum_{i=1}^{I} p_{i,k} \cdot x_{i,k} = 1 \tag{7}$$

$$p_{i,k} \geq 0;\ q_{o,k} \geq 0;\ \forall i, \forall o \tag{8}$$

A complete DEA solves the above linear programming model N times, one for each k. The DMUs having $E_k = 1$ are deemed efficient, whiles those having $E_k < 1$ are deemed inefficient. Thelimitation of this method is that if there are several DMUs having $E_k = 1$, the method cannot provide comparison among these efficient DMUs. To overcome this limitation, Andersen and Petersen (1993) developed procedure for ranking efficient units. The methodology enables an extreme efficient unit to achieve an efficiency score greater than one by removing the k^{th}constraint in the set of constraints given by Eq 6.Moreover, they slightly adjusted the non-negativity constraint in Eq. 8 by imposing the variables $p_{i,k}$ and $q_{o,k}$, to be greater than or equal to a small positive number $\in$. This increases the sensitivity of the result of the DEA analysis to the changes of the levels of the input and the output ($x_{i,n}$ and $y_{o,n}$). The technique is known as the super-efficiency ranking technique and its model is given by Eqs. 9-12. We use this model as the basis for the product cost estimation study proposed in this paper.

Maximize (for a given k):

$$E_k = \sum_{o=1}^{O} q_{o,k} \cdot y_{o,k} \tag{9}$$

subject to:

$$\sum_{o=1}^{O} q_{o,k} \cdot y_{o,n} - \sum_{i=1}^{I} p_{i,k} \cdot x_{i,n} \leq 0;\ \forall n | n \neq k \tag{10}$$

$$\sum_{i=1}^{I} p_{i,k} \cdot x_{i,k} = 1 \tag{11}$$

$$p_{i,k} \geq \in;\ q_{o,k} \geq \in;\ \forall i, \forall o \tag{12}$$

3. Problem Adaptation

In the last decades, several DEA models have been proposed in the operations research and economics literature as tools for the estimation of relative efficiencies and ranking DMUs. In this section we extended the use of DEA beyond this traditional application to cost estimation of products that may be either procured from external suppliers or manufactured in house in built-to-order environment. In such an environment, when an order is placed, only very limited generic design, manufacturing and operational attributes of the products will be known which are referred to as cost drivers. However, it can be assumed that the manufacturer has historical data from similar products with varying degrees of the cost drivers and the costs that have been procured or manufactured in the past. This assumption is in agreement with the vast amount of literature both in parametric and non-parametric cost estimation methods. As it is the case in both parametric and non-parametric cost estimation methods in the literature, we further assume that in the historical data there are N products such that each product has I

number of cost drivers that can be quantified. These cost drivers can be denoted as $\hat{x}_{1,n}, \hat{x}_{2,n}, \cdots, \hat{x}_{I,n}$ for product n. Without loss of generality, in this paper we further assume that the first r cost drivers $(\hat{x}_{1,n}, \hat{x}_{2,n}, \cdots, \hat{x}_{r,n})$ correspond to desirable attributes and the remaining cost drivers $(\hat{x}_{r+1,n}, \hat{x}_{r+2,n}, \cdots, \hat{x}_{I,n})$ correspond to undesirable attributes. By a desirable attribute we mean that, given all other attributes kept unchanged, the more this attribute the product has the better it is (e.g. load carrying capacity) and the opposite applies for undesirable attributes (e.g. weight of sub-assembly). With the above introduction, we present the problem adaptation and the analogy we draw between a DMU and a product for the purpose of using DEA as cost estimation tool.

3.1. *Input adaptation*

For DMU_k shown schematically in Fig. 1(a), given all other things the same, the lesser an input quantity $x_{i,k}$ is the more efficient this DMU. For $\mathrm{PRODUCT}_k$ shown in Fig. 1(b), given all other things the same, the higher desirable attribute $\hat{x}_{i,k}$(for $i \le r$) or the lesser an undesirable attribute $\hat{x}_{i,k}$ (for $i > r$) the better the product is. If we make an analogy between DMU_k and $\mathrm{PRODUCT}_k$, then the inputs to the product should be $x_{i,k} = \frac{1}{\hat{x}_{i,k}}$ for $i \le r$ and $x_{i,k} = \hat{x}_{i,k}$ for $i > r$ so that the lesser an input can be interpreted as the better product as it is the case for DMU_k. This input adaptation is summarized in Eq. (13).

$$x_{i,n} = \begin{cases} \dfrac{1}{\hat{x}_{i,n}} & , \quad \text{for } i = 1, 2, \cdots, r \\ \hat{x}_{i,n} & , \quad \text{for } i = r+1, r+2, \cdots, I \end{cases} \tag{13}$$

3.2. *Output Adaptation*

For DMU_k, given all other inputs and outputs the same, the higher an output $y_{o,k}$ is the more efficient the DMU. For $\mathrm{PRODUCT}_k$, given all other inputs the same, the lower the cost the better the product is. If we make an analogy between DMU_k and $\mathrm{PRODUCT}_k$, then the output from $\mathrm{PRODUCT}_k$ should be $y_{1,k} = 1/\mathrm{cost}$. Thus for this product, given set of inputs, the higher the output (which is now the ratio 1/cost), the better is the product.

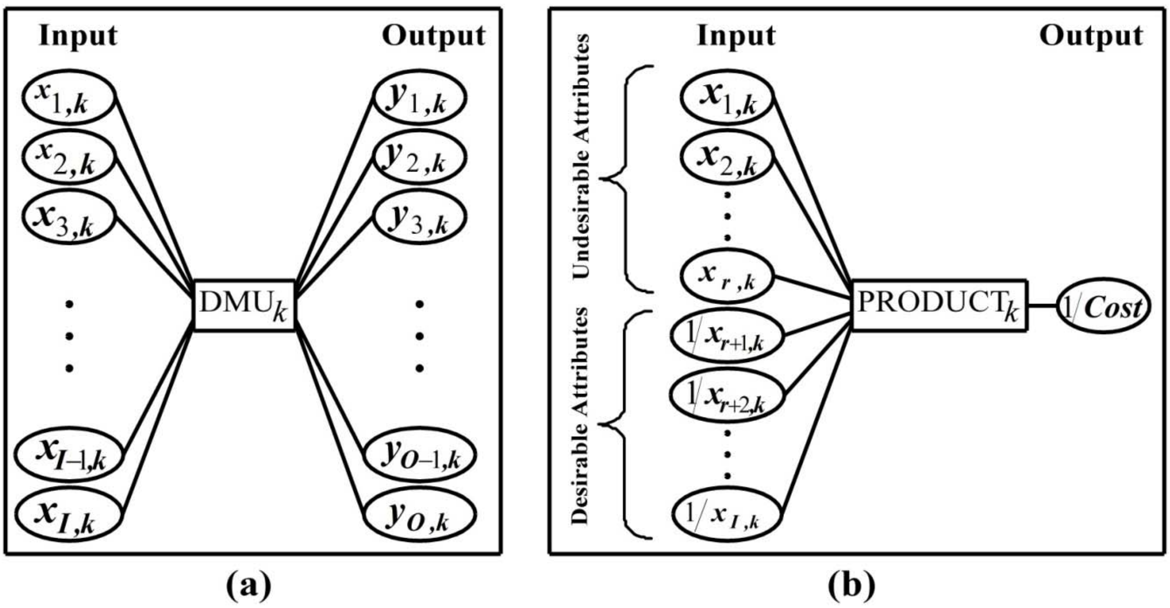


Fig. 1. Analogy between DMU_k and $\mathrm{PRODUCT}_k$ for perfomance evaluation and for cost estimation using DEA

4. Implementation

After problem adaptation, each product n in the historic data has inputs $x_{1,n}, x_{2,n}, \cdots, x_{I,n}$ and an output $y_{1,n} = \frac{1}{\mathrm{Cost}_n}$ having the same interpretation as inputs and outputs of a DMU. Thus, these products in the historic data can be ranked using any DEA model available in the literature. In this paper, we use the super-efficiency DEA model presented in Section 2. Once the products are ranked, the one with the highest $E_k$is considered as the benchmark product. The benchmark can then be used for estimating the cost of new product ($\mathrm{PRODUCT}_{N+1}$) with known inputs $x_{1,N+1}, x_{2,N+1}, \cdots, x_{I,N+1}$and unknown output $y_{1,N+1} = \frac{1}{\mathrm{Cost}_{N+1}}$ where N is the total number of products in the historic data. This cost estimation for product$k = N + 1$ is accomplished by repeatedly solving the super-efficiency DEA model for different trial values of the output $y_{1,N+1}$. If the efficiency of the new product becomes equal to that of the efficiency of the benchmark product for certain trial value of $y_{1,N+1}$, then this trial value is used to estimate the cost of the new product as $\mathrm{Cost}_{N+1} = \frac{1}{y_{1,N+1}}$. The cost found in this way will render the new product as efficient as the benchmark product.

5. Case Study

In this section we present a case study to illustrate the DEA based cost estimation method proposed in this paper. The case study is carried out in collaboration with Bombardier-Aerospace that manufactures and assembles regional and business jets. The products considered for this study are the main landing gears (MLGs) of thirteen different aircraft models assembled in the past by Bombardier Aerospace. A landing gear, also known as an undercarriage, is utilized as an interface between the aircraft and the ground. Fig. 2 below depicts the fully extended MLG of Bombardier Q400 aircraft in flight.

Fig. 2. Main Landing Gear of Bombardier Q400

5.1 Cost drivers and problem adaption

After consulting internal experts, three factors were selected as the potential cost drivers of an MLG. These factors are (1) the maximum takeoff weight of the aircraft (MTOW) on which the MLG is to be fitted, (2) the height and (3) the weight of the MLG. MTOW (which we denote as $\hat{x}_{1,n}$ and measured in pounds) is the heaviest weight at which the aircraft has been shown to meet all the applicable airworthiness requirements. The height of the MLG (denoted as $\hat{x}_{2,n}$ and measured in inches) is the vertical height of the MLG when it is fully extended. The weight of the MLG is denoted as $\hat{x}_{3,n}$ and measured in pounds. The subscript n runs from 1 to $N = 13$ to denote the thirteen landing gears considered in this case study. Once the cost drivers are selected, the first step in applying the proposed cost estimation method is problem adaption. This involves the labeling of each cost driver as a desirable or an undesirable attribute and converting it into an equivalent DMU input by following the procedure outlined in Section 3.

For given height, weight and cost of an MLG, it is obvious that the higher the MTOW is the better the MLG. Therefore, $\hat{x}_{1,n}$ quantifies desirable attribute and thus the corresponding DMU input is $x_{1,n} = \frac{1}{\hat{x}_{1,n}}$. Given MTOW and the weight of an MLG, making it taller requires more sturdy design. It also requires the use of stronger and lighter material to keep the weight unchanged as the height is increased for the same load carrying capacity. Thus, for given MTOW, weight and cost, the taller the MLG is the better it is. This implies that $\hat{x}_{2,n}$ also quantifies desirable attribute and the corresponding DMU input is $x_{2,n} = \frac{1}{\hat{x}_{2,n}}$. The weight $\hat{x}_{3,n}$ of an MLG is an undesirable attribute. For given $\hat{x}_{1,n}, \hat{x}_{2,n}$ and cost of the MLG, the lesser $\hat{x}_{3,n}$ is the better the MLG. The corresponding equivalent DMU input is therefore $x_{3,n} = \hat{x}_{3,n}$. As it was discussed in Section 3, we regard the cost $\hat{y}_{1,n}$ of product (in this case an MLG) as an output and the DMU equivalent output is $y_{1,n} = \frac{1}{\hat{y}_{1,n}}$. Table 1 provides the historic data of thirteen MLGs of different aircraft models and their equivalent DMU input/output pairs. It should be noted that the actual costs are systematically perturbed to protect confidential information.

5.2 *Ranking of MLGs*

In this section, we first illustrate that the efficiency of an equivalent DMU can be used as an indicator of the desirability of product. For this illustration we consider the first landing gear and its equivalent DMU from the historic data given in the first row of Table 1. Fig. 3-shows the efficiency variation of the equivalent DMU as the first input $x_{1,1}$ is increased from smaller value to higher values (where the efficiencies were calculated using the super efficiency DEA model presented in Eqs 9-12). From this figure it can be seen that, as it is the case in any real DMU, the efficiency of the product equivalent DMU falls as the input $x_{1,1}$ is increased while the inputs $x_{2,1}$ and $x_{3,1}$ and the output $y_{1,1}$ are unchanged.

Table 1
Historic data of the thirteen aircraft programs

MLG No. n	MLG Attributes: MTOW $\hat{x}_{1,n}$	Height $\hat{x}_{2,n}$	Weight $\hat{x}_{3,n}$	Cost $\hat{y}_{1,n}$	Equivalent DMU Input: $x_{1,n}$ (10^{-5})	$x_{2,n}$ (10^{-3})	$x_{3,n}$	Output: $y_{1,n}$ (10^{-6})
1	33,000	111	336	63,816	3.0303	9.00	336	15.67
2	36,300	111	336	69,465	2.7548	9.00	336	14.4
3	43,000	112	390	73,794	2.3256	8.97	390	13.55
4	64,500	125	491	125,657	1.5504	8.03	491	7.96
5	37,850	43	200	78,516	2.642	23.10	200	12.74
6	47,600	43	266	117,834	2.1008	23.09	266	8.49
7	53,000	41	329	104,635	1.8868	24.48	329	9.56
8	51,000	42	333	103,552	1.9608	23.81	333	9.66
9	72,750	55	532	103,173	1.3746	18.25	532	9.69
10	80,500	55	532	114,082	1.2422	18.25	532	8.77
11	85,970	55	594	102,595	1.1632	18.18	594	9.75
12	92,500	75	527	104,400	1.0811	13.30	527	9.58
13	98,000	75	527	104,408	1.0204	13.30	527	9.58

Since any increase in $x_{1,1}$ is accompanied by a decrease of the desirable attribute MTOW of the landing gear, decrease in efficiency of the equivalent DMU can be interpreted as decrease in the desirability of the product. Fig. 3-b shows similar results when only $x_{2,1}$ is increased which is again accompanied by decrease in the desirable attribute of the landing gear. Fig. 3-c shows the variation of the efficiency of the equivalent DMU when only $x_{3,1}$ is increased from lower value to higher values. In this case, an increase in the input $x_{3,1}$ is accompanied by an increase in the undesirable attribute which is the weight of the landing gear. Thus, the decrease in the efficiency of the equivalent DMU in Fig. 3 shows the decrease in the desirability of the landing gear because of an increase in an undesirable attribute. Fig. 3-d shows an increase in the efficiency of the equivalent DMU as its output $y_{1,1}$ is increasing. This increase in the efficiently of the equivalent DMU can be interpreted as an increase in the desirability of the landing gear since an increase in$y_{1,1}$ is accompanied by decrease in the cost of the landing gear.

In general, the efficiency of the equivalent DMU of product decreases if there is a decrease in one or more desirable attributes and/or an increase in one or more undesirable attributes of the product. Thus, a group of similar products with varying degrees of desirable and undesirable attributes can be ranked by using efficiencies of their respective equivalent DMUs. Using the super efficiency DEA model presented in Eqs. (9-12) and the data given in the last four columns of Table 1, the efficiencies of the 13 equivalent DMUs of the 13 MLGs were calculated. Table 2 provides the resulting ranking of the MLGs. To gain more insight on this ranking, let us consider some of the landing gears and their ranks.

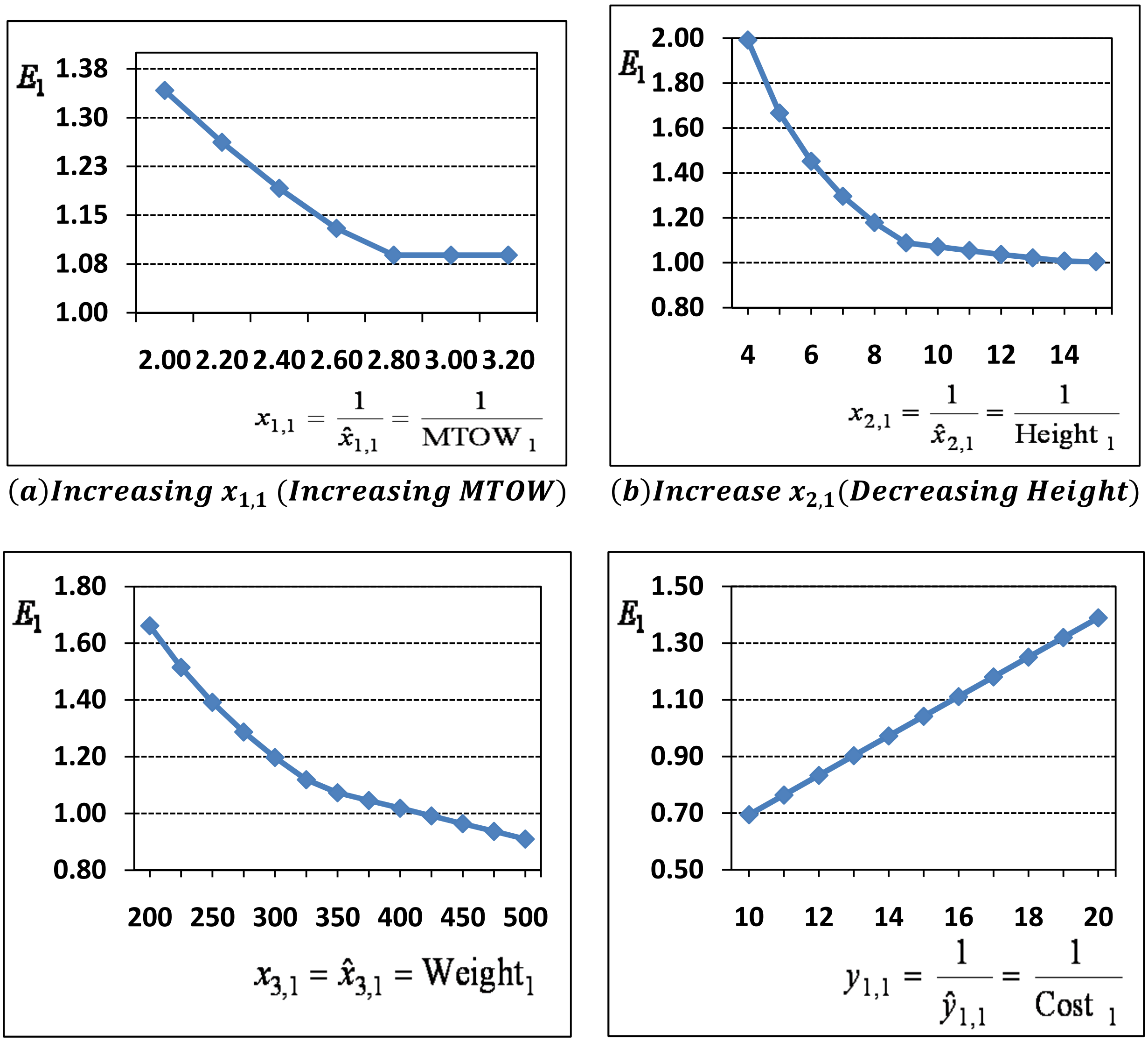
First let us consider landing gears 12 and 13. These two landing gears have identical height and weight and their costs are almost equal. However, the MTOW of landing gear 13 is higher than that of landing gear 12 implying that landing gear 13 is better than landing gear 12. This is clearly indicated by their ranks in Table 2. Now let us consider landing gear 4which has the lowest rank. From Table 1, it can be seen that this landing gear is the most expensive one and its weight is also large. Moreover, its MTOW is smaller compared to landing gears 9 through 13 which do have lower costs than this landing gear. Thus the assignment of the lowest rank to this landing gear by the DEA analysis appeasers to be very reasonable. This ranking of the MLGs illustrates the fact that certain costs in the historic data are too high whereas other costs represent better deals for their corresponding cost drivers. Thus it is important to estimate the cost of a new MLG based on selected benchmark product from the historic data.

Table 2
The ranks of the thirteen MLGs

Rank	1	2	3	4	5	6	7	8	9	10	11	12	13
MLG No.	5	1	13	2	12	3	11	7	9	8	10	6	4
Efficiency	1.367	1.089	1.028	0.978	0.975	0.974	0.858	0.855	0.836	0.833	0.819	0.748	0.673

5.3. Cost Estimation

Once the MLGs are ranked, the cost of new MLG can be estimated based on benchmark MLG as outlined in Section 4. In our example, MLG 5 is the benchmark MLG since it has an equivalent DMU with the highest efficiency, $E_{max} = E_5 = 1.367$. Now, let us assume that we want to estimate the cost of new MLG (denoted hereafter as MLG 14) having its attributes, $\hat{x}_{1,14} = 90{,}000$, $\hat{x}_{2,14} = 75.00$ and $\hat{x}_{3,14} = 400.00$. Let $\hat{y}_{1,14} = \$100{,}000.00$be an initial estimate for the cost of this new MLG. The inputs of the equivalent DMU are therefore,$x_{1,14} = 1.11 \times 10^{-5}$, $x_{2,14} = 1.33 \times 10^{-2}$ and $x_{3,14} = 400.00$, and its output $y_{1,14} = 1 \times 10^{-5}$. Using this input/output information along with that of the thirteen other equivalent DMUs from the historicdata, the efficiency E_{14} of DMU 14 is determined by solving the super efficiency DEA model (Eq. (9) to Eq.(12)) for $k = 14$ and its value is 1.155. However, this efficiency is lower than the benchmark, $E_{max} = 1.367$. This implies that the initial estimate of the cost of the new MLG, $\hat{y}_{1,14} = \$100{,}000.00$, is too high for DMU 14 to be as efficient as the benchmark. Thus we have to try a lower cost estimate so that E_{14}can be as close as E_{max}. Fig. 4 shows the efficiency of DMU 14 at several different estimates of the cost of MLG 14. In this figure, it can be seen that for the DMU 14 to have the benchmark efficiency, E_{max} = 1.367, the cost of MLG 14 should be as low as \$84,500.00. This implies that at historically observed highest efficiency, MLG 14 can be manufactured with that low cost. The cost of this new MLG based on the average historical efficiency ($E_{average}$ = 0.962) is \$124,820.00. Thus the manufacturer (or the supplier) has to strive to produce this landing gear with cost lower than \$124,820.00 to stay at higher level of efficiency than the historical average efficiency. At the lowest historical efficiency, E_{min} = 0.673, the cost of MLG 14 could be as high as \$171,600.00.



(a) ***Increasing*** $x_{1,1}$ ***(Increasing MTOW)*** (b) ***Increase*** $x_{2,1}$ ***(Decreasing Height)***

(c) ***Increasing*** $x_{3,1}$ ***(Increasing Weight)*** (d) ***Increasing*** $y_{1,1}$ ***(Decreasing Cost)***

Fig. 3. Illustration of the effect of increasing an input or an output on the efficiency of an equivalent DMU of a landing gear.

6. Relationships with other Methods

Cost estimation methods based on parametric or non-parametric approaches assume that the historical data is a true representation of the relationship between the cost drivers and the corresponding costs. This implies that all the products in the historic data are considered to be equally good in terms of their costs. Given this fact, we foresee that the cost estimates using these methods will be close to those that can be obtained using the purposed DEA based method if the cost estimation using the DEA method is based on the average observed efficiency. To illustrate this empirical relationship that may exist between parametric or non-parametric methods and the DEA based approach, we consider cost estimation of six new MLGs where attributes data was arbitrarily generated (see Table 3).

For this illustration, we choose the back propagation ANN non-parametric cost estimation methods as ANN has been proven to be superior compared to parametric and other nonparametric methods (Cavalieri et al., 2004; Caputo and Pelagagge, 2008; Verlinden et al., 2008; Dura et al., 2009). The ANN used in this work has been detailed in the Appendix and was trained using the thirteen data points given in Table 1. Fig. 5 provides the cost estimates of the six new MLGs using the ANN and the DEA method based on the historic average and maximum efficiencies. In this figure it can be seen that when the average historic efficiency is used as a reference for cost estimation using the DEA method, the cost estimates found are very close to those found using the ANN. This illustrate the empirical relationship that we foresee to exist between existing methods and the DEA based method proposed in this paper. When the historically observed highest efficiencies are used as a reference, the proposed method provides the lowest possible cost that can be incurred in manufacturing the products if the manufacturer operates at its historically observed highest efficiency

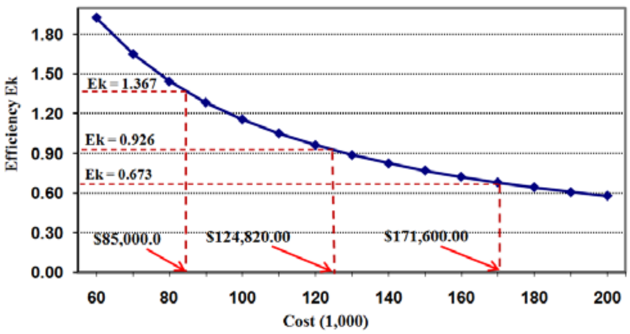


Fig.4. E_k values of the new MLG at different trial cost values and its cost estimates corresponding to the historically observed maximum, average and minimum efficiencies

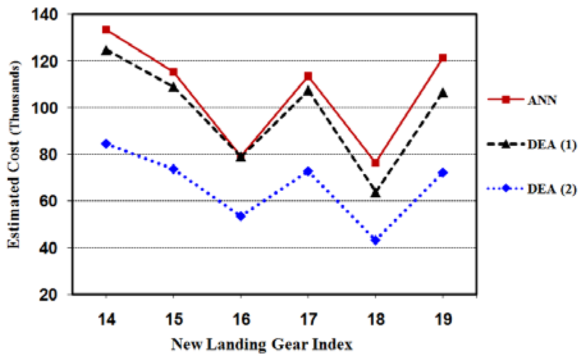


Fig. 5. Estimation using DEA and ANN models

Table 3
Arbitrarily generated data for six new MLGs

MLG No.	MTOW $\hat{x}_{1,n}$	Height $\hat{x}_{2,n}$	Weight $\hat{x}_{3,n}$
14	90,000	75	400
15	60,000	40	300
16	55,000	75	500
17	50,000	110	200
18	30,000	80	350
19	85,000	55	520

7. Discussion and conclusions

Parametric and non-parametric cost estimation methods assume that the historic data set is a true representation of the relationship between the cost drivers and the cost. This is evidenced by the fact that the objective function used in the search of parameters or connection weight is the goodness of the fit of the input/output from these methods to that of the historic data. However, because of efficiency variations of the manufacturers and suppliers, changes in suppliers, market fluctuations, and several other reasons, certain costs in the historic data may be too high whereas other costs may represent better deals for their corresponding cost drivers. This was clearly demonstrated by the ranking of the products using the DEA method. In such scenarios, it may be important to base the cost estimation on selected benchmarks. The proposed method provides cost estimations based on the observed highest efficiency or an efficiency level at which the manufacturer is willing to operate. In addition to providing such valuable information, the proposed method can also be used to generate cost estimates that are close to those that can be obtained using parametric or non-parametric methods if the cost estimation using the proposed method is based on the observed average efficiency. Another added advantage of the proposed method is that it can also be used to rank competitive brands and models of products available in the market and help purchasers in performing guided selection of these products.

References

Adler, N., Friedman, L., & Sinuany-Stern, Z. (2002). Review of ranking methods in the data envelopment analysis context. *European Journal of Operational Research*, 140, 249–265.

Andersen, P., & Petersen, N. (1993). A procedure for ranking efficient units in data envelopment analysis. *Management Science*, 39, 1261–1294.

Bode, J. (1998). Decision support with neural networks in the management of research and development: Concepts and applications to cost estimation. *Information & Management*, 34, 33–40.

Bode, J. (2000). Neural networks for cost estimation: Simulation and pilot applications. *International Journal of Production Research*, 38, 1231–1254.

Caputo, A. C., & Pelagagge, P. M. (2008). Parametric and neural methods for cost estimation of process vessels. *International Journal of Production Economics*, 112, 934–954.

Cavalieri, S., Maccarrone, P., & Pinto, R. (2004). Parametric vs. neural network models for the estimation of production costs: A case study in the automotive industry. *International Journal of Production Economics*, 91, 165–177.

Chan, K. Y., Kwong, C. K., & Tsim, Y. C. (2010). A genetic programming based fuzzy regression approach to modeling manufacturing processes. *International Journal of Production Research*, 48, 1967–1982.

Charnes, A., Cooper, W., & Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operations Research*, 2, 429–444.

Dean, E. P. (2005). Parametric cost deployment. *Proceedings of the Seventh Symposium on Quality Function Deployment.*USA, 27–34.

Dura, O., Rodriguez, N., & Consalter, L. A. (2009). Neural networks for cost estimation of shell and tube heat exchangers. *Expert Systems with Applications*, 36, 7435–7440.

Foussier, P. M. M. (2006). *From Product Description to Cost: A Practical Approach. Vol. 2, Building a Specific Model*. Berlin: Springer.

Layer, A., Brine, E. T., Van, H. F., Kals, H., & Haasis, S.(2002). Recent and future trends in cost estimation.*International Journalof Computer Integrated Manufacturing*, 15, 499–510.

Pandya, A. S., & Macy, R. B.(1996).*Pattern Recognition with Neural Networks in C++*.Florida: CRC Press.

Phaobunjong, K. & Popescu, C. M. (2003). Parametric cost estimating model for buildings. AACE International Transactions.

Seo, K. K., Park, J. H., Jang, D. S., & Wallace, D. (2002). Prediction of the life cycle cost using statistical and artificial neural network methods in conceptual product design. *Journal of Computer Integrated Manufacturing*, 15, 541–554.

Verlinden, B., Duflou, J. R., Collin, P., & Cattrysse, D.(2008). Cost estimation for sheet metal parts using multiple regression and artificial neural networks: A case study. *International Journal of Production Economics*, 111, 484–492.

Zhang, Y. F., Fuh, J. Y., & Chan, W. T.(1996).Feature-based cost estimation for packaging products using neural networks. *Computers in Industry*, 32, 95–113.

Zhang, Y. F. & Fuh, J. Y. H.(1998). A neural network approach for early cost estimation of packaging products. *Computers & Industrial Engineering*, 34, 433–450.

Zheng, G. & Zhang, P.(2006).Meta-heuristic algorithms for parameter estimation of semi-parametric linear regression models. *Computational Statistics & Data Analysis*, 51, 801–808.

Appendix: Artificial Neural Network

In this appendix, we provide some details of the back propagation ANN used to predict the costs of the six new MLGs. The architecture of this ANN is given in Fig. 6. It was trained using the thirteen historic data points from the existing programs. For the purpose of training and predication, the data points were normalized using Eqs. (14) and (15) for the inputs and outputs, respectively. The input vector to this neural network comprises the three normalized attributes of a landing gear. The input layer passes the input vector to the hidden layer through the connections. The output of a neuron out_iin the hidden and output layers is calculated using a sigmoid function given in Eq. 16 where net_i is weighted input to this neuron and Q is a shape parameter was set at 0.9. A bias neuron does not have an input and its output is 1.0. The connection weights of the ANN after training are given in Table 4. Details of this type of neural network and its training algorithm can be found in Pandya and Macy (1996).

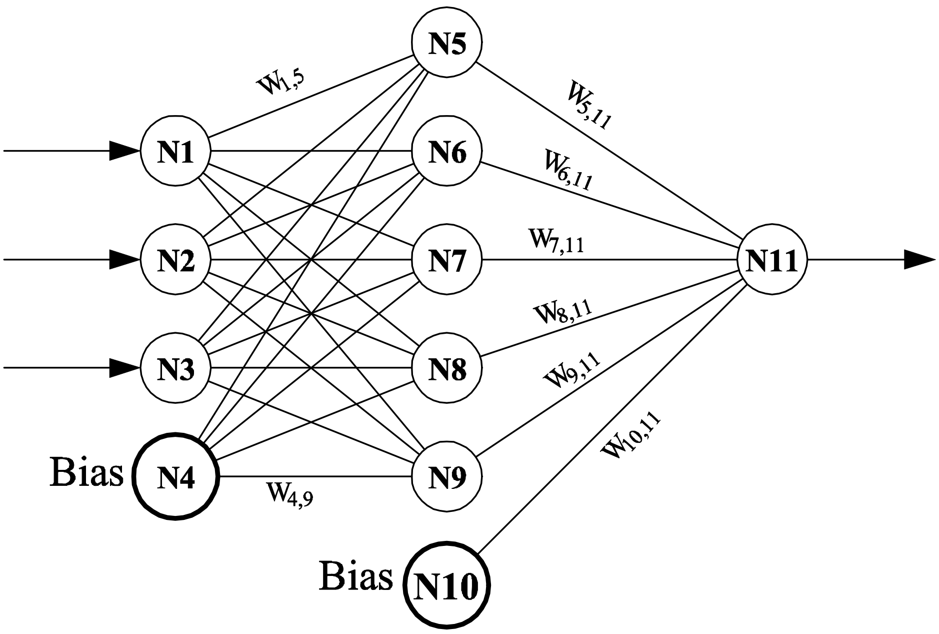


Fig. 6. The artificial neural network used for cost estimation

$$\text{normilized}\left(\hat{x}_{i,j}\right) = \frac{\hat{x}_{i,j} - 0.8 \times \min_{i=1}^{i=13}(\hat{x}_{i,j})}{1.2 \times \max_{i=1}^{i=13}\left(\hat{x}_{i,j}\right) - 0.8 \times \min_{i=1}^{i=13}(\hat{x}_{i,j})} \quad (14)$$

$$\text{normilized}\left(\hat{y}_{i,j}\right) = \frac{\hat{y}_{i,j} - 0.8 \times \min_{i=1}^{i=13}(\hat{y}_{i,j})}{1.2 \times \max_{i=1}^{i=13}\left(\hat{y}_{i,j}\right) - 0.8 \times \min_{i=1}^{i=13}(\hat{y}_{i,j})} \quad (15)$$

$$out_i = f(net_i) = \frac{1}{1 + e^{\frac{-net_i}{Q}}} \quad (16)$$

Table 4
Weights for the artificial neural network shown in Fig. 6

	Hidden Layer Weights $W_{i,j}$					Output Layer Weights	
	j						
i	5	6	7	8	9	i	$W_{i,11}$
1	2.1632	4.5333	1.2724	10.8861	0.3612	5	13.1887
2	5.3344	-1.9746	2.7691	12.7256	4.0436	6	-5.2210
3	-0.6848	7.4131	-1.2427	-0.6552	-8.4051	7	-0.0567
4	-1.7799	-3.5752	2.1494	-6.8283	1.1156	8	-4.8774
						9	-6.1845
						10	0.7127