

Contents lists available at GrowingScience

## Decision Science Letters

homepage: [www.GrowingScience.com/dsl](http://www.GrowingScience.com/dsl)**A case study to estimate costs using neural networks and regression based models****Adil Salam<sup>a\*</sup>, Fantahun M. Defersha<sup>b</sup> and Nadia Bhuiyan<sup>a</sup>**<sup>a</sup>*Department of Mechanical and Industrial Engineering, Concordia University 1455 de Maisonneuve W., Montreal, Quebec, Canada, H3G 1M8*<sup>b</sup>*School of Engineering, University of Guelph, 50 Stone Road East, Guelph, Ontario, Canada, N1G 2W1***ARTICLE INFO***Article history:*

Received September 9, 2010  
Received in revised format June 5, 2012  
Accepted 11 July 2012  
Available online  
July 11 2012

*Keywords:*

*Target cost*  
*Parametric models*  
*Neural networks*

**ABSTRACT**

Bombardier Aerospace's high performance aircrafts and services set the utmost standard for the Aerospace industry. A case study in collaboration with Bombardier Aerospace is conducted in order to estimate the target cost of a landing gear. More precisely, the study uses both parametric model and neural network models to estimate the cost of main landing gears, a major aircraft commodity. A comparative analysis between the parametric based model and those upon neural networks model will be considered in order to determine the most accurate method to predict the cost of a main landing gear. Several trials are presented for the design and use of the neural network model. The analysis for the case under study shows the flexibility in the design of the neural network model. Furthermore, the performance of the neural network model is deemed superior to the parametric models for this case study.

© 2012 Growing Science Ltd. All rights reserved.

**1. Introduction**

Will this be a feasible project? Do we know how much it will cost? Are we able to afford this? These are a series of questions on the minds of most people and especially management when they are considering either introducing a new product to the market, or even if they are considering purchasing a product. To have the ability to accurately predict any cost in the aerospace industry is of utmost importance in order to gain an advantage over its competitors and suppliers. The importance of having this capability is especially vital in the aerospace industry, due to the fact that some of those products cost several million dollars per ship set. Several case studies have been conducted showing the costs of products and projects are not properly allocated and tend to go over budget (Assaf & Al-Hejji; 2006; Frimpong, 2000; Bounds, 1998; Murmann, 1994; Norris, 1971).

Furthermore, determining target costs will have other benefits with respect to forecasting, scheduling, bidding, fact based negotiation and in making strategic decisions. This paper presents a case study, in

\* Corresponding author. Tel. +1 (514) 631-6775 ext. 1240  
E-mail addresses: [adil.salam@novartis.com](mailto:adil.salam@novartis.com) (A. Salam)

the form of an empirical study, carried out in collaboration with Bombardier Aerospace (BA). BA is well known as a global leader in the manufacturing and assembly of regional and business jets.

There are a various numbers of tools, models and methodologies available in the cost estimation industry. A bottom up approach entails estimating the cost of each and every activity according to their manufacturing sequence, summing these activities and finally allocating appropriate overheads. On the other hand, the estimation techniques that are under study, more particularly parametric and neural networks (NN) are considered a top down approach .

A top down approach takes into consideration the total effort of the project as a whole and assessing the most significant cost estimating relationship (CER), otherwise known as the target cost model. The two models previously mentioned will be developed and compared in order to estimate the target cost of a main landing gear (MLG) at BA. Authors such as Colmer et al. (1999), Bounds (1998) and Norris (1971) have emphasized on the importance of understanding the amount of resources, time, and cost required to complete a product. Even though this empirical study was a particular case at BA, the concepts and models used to estimate the cost of the MLG presented in this paper can be applied in any particular industry. The mathematical models presented in this paper can be used for a wide range of applications, such as cost estimations, make/buy decisions, project feasibility, and profitability for the company.

The rest of the paper is organized as follows; Section 2 presents the methodology utilized. It also gives an overview of the component under study, the MLG. Section 3 introduces both the parametric and neural network models developed for predicting the target costs. It also discusses the prominent factors selected as the cost drivers. The development of both types of models along with a detailed analysis is presented in Section 4. Finally, the conclusions and suggestions for future research highlighted in Section 5.

## **2. Methodology**

BA is interested in estimating the target costs of various commodities and systems for their aircrafts. Even though several preliminary models have been created and some are in the process of being developed, this paper shows a regression based parametric model and a NN model developed to estimate the costs of the MLGs, which is a new approach at BA.

### *2.1 Landing Gear (LG)*

The undercarriage of an aircraft, commonly referred to as the landing gear, is utilized as an interface between the aircraft and the ground. The complete landing gear assembly can be divided into sub-systems. The main sub-systems are the MLG, nose landing gear, extension and retraction system, alternate release system, steering system and brake control system. The landing gear at BA is a major commodity in which absorbs roughly 5-10% of the overall cost of the aircraft. Therefore, to accurately predict its cost can serve as a great benefit. The focus of this paper is to develop target cost models for MLGs. The main functions of the landing gear are to absorb loads upon landing, taxiing and braking. The load during landing is absorbed by the gears and it is proportional to the maximum takeoff weight (MTOW). The MTOW, which is typically measured in the unit of pounds (lbs.), is the maximum allowable weight for an aircraft to take-off to meet the requirements of airworthiness authorities.

## **3. Target cost estimation**

As previously mentioned, a comparative study was conducted for a regression based parametric model to a neural networks model. The basic theory of both methods is discussed in the section below. Both of these methods are used to predict the target costs of MLGs based upon certain input variables that were considered as cost drivers.

### 3.1. Cost drivers

Upon researching, analyzing the technical specifications and speaking to experts in the field of landing gears, the following three factors, which are provided to a potential supplier to make a bid, were selected as the CER to estimate the target cost:

1. Weight of the MLG
2. Maximum takeoff weight (MTOW)
3. Height of the MLG

The weight of the MLG is measured in pounds. The MTOW has already been discussed in the previous section, and the height of the MLG, measured in inches, is the vertical height of the MLG when it is fully extended.

### 3.2. Regression Based Parametric Models

When a desired output,  $y$  depends on several input variables, as is the case in the problem under study, the multiple linear regression model (MLRM) can be utilized (Kutner et al., 2004). The basic equation of the MLRM is as follows:

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_j X_j + \varepsilon \quad (1)$$

where,

$y$ , response dependent output on  $k$  predictor values

$\beta_j$ , regression coefficient

$X_j$ ,  $j^{\text{th}}$  independent variable

$\varepsilon$ , is the error

The parametric model shown below has been used recently in another similar application, to estimate the required design effort to complete a project had promising results (Salam *et al.*, 2009, 2008; Bashir & Thomson, 2004). Furthermore, the generic model shown below has been previously validated with empirical evidence of Boehm (1991), and Walston and Felix (1977).

$$\hat{y} = \beta_0 X_1^{\beta_1} X_2^{\beta_2} \dots X_m^{\beta_m} \quad (2)$$

where  $\hat{y}$  is estimated target cost,  $X_m$ ,  $m^{\text{th}}$  is cost driver and  $\beta_m$ , constant (weight) estimated from historical data. It should be noted that Eq. (2) is not a linear equation and is not in the suitable form of a MLRM. In order to overcome this problem, if the natural log (ln) is taken of both sides the equation will have the form shown in Eq. (3).

$$\ln(\hat{y}) = \ln(\beta_0) + \beta_1 \ln(X_1) + \beta_2 \ln(X_2) + \dots + \beta_m \ln(X_m) \quad (3)$$

Now in its current form it is clear that Eq. (3) is in the proper form in order to carry out the regression, provided the natural log of the input variables are also taken. However, as the data set is small, the jackknife technique for regression models is applied to ameliorate the results.

#### 3.2.1 Jackknife technique

The jackknife technique is used to determine the regression coefficients of each of the model parameters. This technique was originally a computer-based method for estimating biases and the standard errors. According to Efron and Tibshirani (1993), this technique is commonly used not only

to improve the problem of biased estimation due to small sample size, but also in situations where the distribution of the data is hard to analyze. In this technique, the data are divided into sub-samples, and the sub-samples are obtained by deleting one observation at a time. The calculations are carried out for each sub sample. Given a data set  $x = (x_1, x_2, x_3, x_n)$ , the  $i^{\text{th}}$  jackknife sample  $x_i$  is defined to be  $x$  with the  $i^{\text{th}}$  data point removed. The pseudo-values,  $Ps_i$ , are determined using the following equation:

$$Ps_i = ns\beta - (ns-1)\beta_{-i} \quad (4)$$

where,

$Ps_i$ , Pseudo-value for the entire sample, omitting sub-sample  $i$

$ns$ , Number of sub-samples

$\beta$ , Least-squares estimator of the whole sample

$\beta_{-i}$ , Least-squares estimator for the entire sample, omitting sub sample  $i$

The jackknife estimator  $\tilde{\beta}$  is determined as follows:

$$\tilde{\beta} = \frac{\sum_{i=1}^{ns} Ps_i}{ns} \quad (5)$$

### 3.2 Complex Parametric Models

The CNLM used in this research has the following notation.

$$\hat{y} = a_0x_1^{b_0} + a_1x_2^{b_1} + a_2x_3^{b_2} + c_0x_1^{d_0}x_2^{d_1} + c_1x_1^{d_2}x_3^{d_3} + c_2x_2^{d_4}x_3^{d_5} + e_0x_1^{f_0}x_2^{f_1}x_3^{f_2} \quad (6)$$

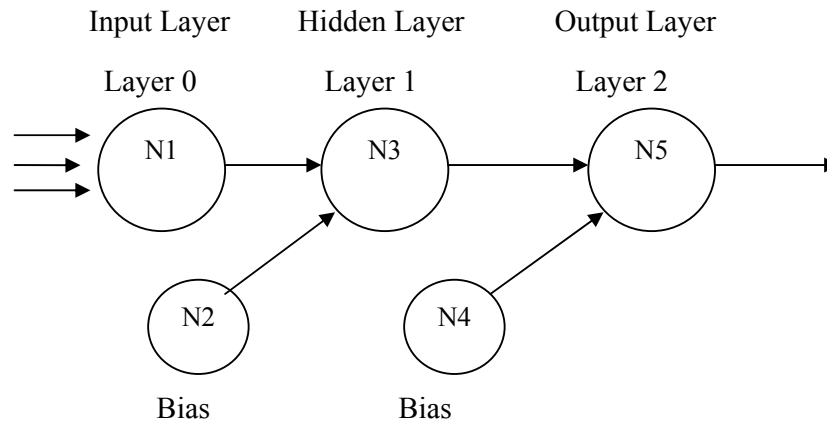
As was the case in the regression models, the terms  $x_i$  represent the cost drivers, and remaining terms are the constants. As this equation is not in the form of a regression model, the constants will have to be determined analytically. The manner in which the constants will be determined will be using the gradient descent algorithm (GDA). The GDA is an optimization tool to find the local minima of a function (Snyman, 2005). In order to determine the constants, the function to be minimized is the square error of the predicted versus the actual costs (ie.  $\sum(y - \hat{y})^2$ ).

The gradient,  $\nabla$  for each of the constants will calculate the amount the constant has to be changed (ie. delta) in order to minimize the function, the square error. Furthermore, the value of the constant will be adjusted by multiplying it by a step rate,  $\eta$ . Each of the constants will be adjusted each iteration until the specified stopping criterion (ie. acceptable change in error) is fulfilled.

Parametric CERs depend on a pre-determined cost function to establish the target cost. However, using non-parametric models such as the models based upon artificial intelligence, such as neural networks (NN) does not require pre-determined relationships. NN models have the ability of self-determining relationships of the variables to predict the cost. Therefore, the next model developed is based upon NN, and is discussed in the following section.

### 3.3 Neural Networks

Multilayer NNs which are trained by back propagation, explained below, are amongst the most popular forms of NNs. They are able to classify data and approximate functions based on a set of sample data, known as the training data. In literature, it is shown that an NN with a single hidden layer and a non-linear activation function can approximate the decision boundaries of a wide range of complexity (Kecman, 2001). This property is used to investigate the applicability of neural networks for cost estimation of the MLGs for BA. The NN used in this study is depicted in Fig. 1.



**Fig. 1.** The single hidden layer neural network used in this study

An input vector representing the three cost drivers (see Section 3.1) is incident on the input layer. It is then distributed to the hidden layer and finally to the output layer via its weighted connections. As can be seen in the figure a bias neuron is also present. The bias neuron always outputs a unity without having an input. The bias neuron is very important, since the error-back propagation neural network without bias neuron for hidden layer will not learn (Kecman, 2001). Besides the bias neuron, each of the remaining neurons in the network operates by taking the sum of its weighted inputs and passing the result through a nonlinear activation function. This is shown mathematically in Eq. (6).

$$out_{i,l} = f(net_i) = f\left(\sum_j w_{j,i,l} out_{j,l-1}\right) \quad (6)$$

where

$out_{i,l}$  output of  $i^{\text{th}}$  neuron in layer  $l$

$w_{j,i,l}$  weight for the connection between  $j^{\text{th}}$  neuron of layer  $l-1$  and  $i^{\text{th}}$  neuron of layer  $l$

$f$ , nonlinear activation function.

There are several conventionally used choices for this activation function. One of the most frequently used functions (Pandya & Macy, 1996), also used in this study, is the sigmoid function given in Eq. (7).

$$f(net_i) = \frac{1}{1 + e^{-\frac{net_i}{Q}}} \quad (7)$$

The computational simplicity of the derivative of this function simplifies the formulation of the equations needed for the training process. Moreover, this function is bounded between  $\pm 1$  ensuring that certain signals remain within a range and introduces non-linearity to the model. The term  $Q$  in this equation is referred to as the temperature of the neuron and it determines the shape of the sigmoid. It is used to tune the network in order to improve its convergence behavior.

For the NN to be able to predict a cost given the three cost drivers as inputs, it must first be trained using a training set. The training is the process of a successive and a systematic adjustment of the weights in order to minimize a defined measure of error, which is typically the square error. In this process, each time an input vector  $p$  (the three cost drivers) from the training set is presented to the network, the difference between the desired and the actual output is computed and the each weight is adjusted by an amount  $\Delta w_{j,i,l}$ , given in Eq. (8).

$$\Delta w_{j,i,l} = \eta \times \delta_{p,i,l} \times O_{p,i,l} \quad (8)$$

where  $\eta$  is learning rate,  $O_{p,j,l}$  is output of the  $j^{\text{th}}$  neuron of layer  $l$  and  $\delta_{p,j,l}$  error signal at the  $j^{\text{th}}$  neuron of layer  $l$ .

The error signal is computed using Eq. (9) for the neuron in the output layer.

$$\delta_{p,1,2} = (t_{p,1,2} - O_{p,1,2}) \times O_{p,1,2} \times (1 - O_{p,1,2}) \quad (9)$$

It should be noted that  $t_{p,1,2}$  is the desired value at the neuron in the output layer corresponding to the  $p^{\text{th}}$  presentation of the training set vector. The training algorithm will try to obtain the output value as close as possible to  $t_{p,1,2}$  while minimizing the error. Eq. (10) is used to compute the error signal for the neurons in the hidden layer.

$$\delta_{p,j,1} = O_{p,j,1} \times (1 - O_{p,j,1}) \times \delta_{p,1,2} \times w_{j,1,2} \quad (10)$$

This training algorithm is known as training by back-propagation, and the complete derivation of the algorithm can be seen in Pandya and Macy (1996).

#### 4. Data Analysis

In order to conduct this empirical study using both the parametric and the NN model, historical data is required. Table 1 below shows the costs of thirteen MLGs along with the corresponding technical data for the three cost drivers for thirteen types of aircraft programs at BA.

**Table 1**  
Historical data

Program	X <sub>1</sub> : Weight	X <sub>2</sub> : MTOW	X <sub>3</sub> : Height	y <sub>i</sub> : Masked Cost
1	335.77	33,000	111.16	63,816
2	335.77	36,300	111.16	69,465
3	389.97	43,000	111.54	73,794
4	490.58	64,500	124.58	125,657
5	199.58	37,850	43.29	78,516
6	265.62	47,600	43.30	117,834
7	328.81	53,000	40.85	104,635
8	333.00	51,000	42.00	103,552
9	532.00	72,750	54.80	103,173
10	532.00	80,500	54.80	114,082
11	594.00	85,970	55.00	102,595
12	526.97	92,500	75.17	104,400
13	526.97	98,000	75.17	104,408

It should be noted that the “actual” costs are not revealed and are masked to protect confidential information. The manner the data is masked is important. The characteristics of the original data should be maintained, thus a masking technique developed in presented in Muralidhar *et al.* (1999) is utilized, and Bombardiers permission was obtained to use the masked data. For both of the types of models, sub-samples of the data were randomly generated, where ten programs were used to generate or train the model, and the remaining three were used for validation purposes. Three trials were conducted for this analysis. For the purpose of validation, the programs (6, 7, 13), (3, 7, 9), and (2, 8, 12) were removed from the data for Trials 1, 2, and 3, respectively.

##### 4.1 Analysis Based on Regression

For each trial a MLRM was carried out and it resulted in an equation to estimate the target cost of MLGs. The errors of the testing and validation data is obtained and will be further discussed in Section 4.3. Furthermore, with the MLRM, an analysis of variance (ANOVA) is conducted to see if the factors considered are significant. The confidence interval for significance is set at 90%. The

procedure highlighted by Kutner *et al.* (2004) is used to determine the significant factors. If factors are considered insignificant the analysis is repeated by removing one factor at a time and the analysis is continuously repeated until only significant factors remain. The average generated equations to estimate the target costs are summarized in Table 2 below. In all three of the trials there is only 1 factor, MTOW which is rendered as the significant factor.

**Table 2**

Regression based parametric models for Trial 1

$$\ln \hat{y}_{3, RM} = 5.16 - 0.25 \ln x_1 + 0.70 \ln x_2 + 0.03 \ln x_3$$

$$\ln \hat{y}_{2, RM} = 5.57 - 0.19 \ln x_1 + 0.64 \ln x_2$$

$$\ln \hat{y}_{1, RM} = 5.88 + 0.51 \ln x_2$$

#### 4.2. Analysis Based on Complex Parametric Models

For modeling purposes, the values of input and cost data were normalized to facilitate in the convergence of the model.

$$x_1 \rightarrow x_{1(original)}/1,000$$

$$x_2 \rightarrow x_{2(original)}/100,000$$

$$x_3 \rightarrow x_{3(original)}/1,000$$

$$y \rightarrow y_{(original)}/1,000,000$$

As mentioned in Section 3.2, the gradient descent algorithm (GDA) was used to determine the coefficients of the model. For all 3 Trials, using the GDA, the models converged and resulted in the following equations.

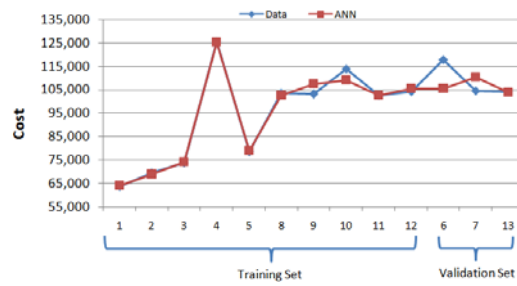
$$\hat{y}_{CNLM, Trial 1} = 0.053x_1^{0.664} + 0.189x_2^{0.464} - 0.314x_3^{0.590} - 0.176x_1^{0.605}x_2^{0.678} + 0.179x_1^{0.623}x_3^{0.728} + 0.180x_2^{0.592}x_3^{0.725} + 0.331x_1^{0.574}x_2^{0.544}x_3^{0.697} \quad (11)$$

$$\hat{y}_{CNLM, Trial 2} = 0.104x_1^{0.682} + 0.304x_2^{0.523} - 0.428x_3^{0.408} - 0.339x_1^{0.671}x_2^{0.913} + 0.242x_1^{0.623}x_3^{0.722} + 0.254x_2^{0.584}x_3^{0.714} + 0.402x_1^{0.560}x_2^{0.521}x_3^{0.684} \quad (12)$$

$$\hat{y}_{CNLM, Trial 3} = 0.139x_1^{0.693} + 0.257x_2^{0.516} - 0.408x_3^{0.453} - 0.319x_1^{0.664}x_2^{0.875} + 0.223x_1^{0.622}x_3^{0.718} + 0.208x_2^{0.574}x_3^{0.715} + 0.364x_1^{0.565}x_2^{0.520}x_3^{0.688} \quad (13)$$

#### 4.3. Analysis Based on Neural Networks

The NN analysis is carried out for all the 3 trials as the linear regression. For each of the trials the number of neurons in the single hidden layer is set at 5. The temperature  $Q$  is set at 0.9, and the learning rate  $\eta$  is set at 0.1. The stopping criterion,  $\delta$  is the maximum value of acceptable error for the normalized training data. The value was determined by adjusting for each of the trials in order to minimize the maximum error of the testing and validation errors, the final selected value was 0.05. Furthermore, the maximum number of allowable iterations was set at 192,000,000 to prevent the model from “crashing” if no convergence is obtained. The value of delta was set at Fig. 2 depicts the corresponding plot of the data showing the actual (Data) versus the prediction (NN) for both the training and validation points for Trail 1. In the plot shown in Fig. 2, we can see that the NN is able to make a good prediction both on the training and the validation points. Such interesting results are also obtained in the others trials. This shows that the NN is able to generalize even after being trained with the limited data sets available in this problem. The comparison between the parametric model and the NN for all the trials is summarized in the following section.



**Fig. 2.** Plot of the masked cost versus the prediction using NN

#### 4.4 Summary and Comparative Analysis

Tables 5-7 below summarize the results for the regression, the complex parametric and the NN models for Trials 1-3, respectively. It tabulates the maximum error for all the trials on the training, validation and all the combined data.

**Table 5**

Comparison of Errors for Trial 1

Program	RM (%)	CPM (%)	NN (%)
1	10.10	7.01	0.50
2	6.16	4.75	0.96
3	8.89	12.33	0.38
4	21.45	14.79	0.21
5	4.07	3.09	0.63
8	28.20	10.32	0.84
9	14.61	0.07	4.25
10	15.39	6.51	4.26
11	1.69	4.35	0.01
12	3.19	11.56	1.13
<b>6</b>	<b>28.20</b>	<b>23.08</b>	<b>10.33</b>
<b>7</b>	<b>14.61</b>	<b>9.61</b>	<b>5.65</b>
<b>13</b>	<b>16.87</b>	<b>13.95</b>	<b>0.51</b>
Training	11.38	7.48	1.32
Validation	19.89	15.55	5.50

**Table 6**

Comparison of Errors for Trial 2

Program	RM (%)	CPM (%)	NN (%)
1	19.76	2.02	0.11
3	15.71	2.75	0.05
4	19.46	15.55	0.00
5	3.10	11.29	0.19
6	24.38	10.47	0.18
7	10.92	1.20	0.31
9	3.16	3.27	2.04
10	2.66	5.49	4.30
11	11.26	0.14	1.84
13	15.49	8.10	0.60
<b>2</b>	<b>14.51</b>	<b>13.69</b>	<b>4.93</b>
<b>8</b>	<b>11.43</b>	<b>2.52</b>	<b>1.72</b>
<b>12</b>	<b>12.74</b>	<b>8.59</b>	<b>0.81</b>
Training	12.59	6.03	0.96
Validation	12.89	8.27	2.49

**Table 7**

Comparison of Errors for Trial 3

Program	RM (%)	CPM (%)	NN (%)
1	20.40	11.28	0.01
2	22.86	9.05	0.24
4	20.23	19.81	0.18
5	3.33	12.61	0.32
6	24.59	10.88	0.34
8	10.46	1.85	0.38
10	7.83	5.45	4.91
11	6.68	4.74	4.56
12	7.61	2.19	0.74
13	13.34	5.48	0.40
<b>3</b>	<b>8.13</b>	<b>14.33</b>	<b>10.77</b>
<b>7</b>	<b>12.73</b>	<b>2.45</b>	<b>7.61</b>
<b>9</b>	<b>12.10</b>	<b>5.17</b>	<b>5.96</b>
Training	13.73	8.33	1.21
Validation	10.99	7.32	8.11

As can be seen from Table 5-8, the regression model is the worst in terms of validation error for all 3 Trials. The NN based model outperforms both former models in all 3 Trials in terms of error on the training data, and is superior in the majority of the cases when the validation errors are compared. Furthermore, a *t-test* comparing the validation errors for the 3 models infers that the NN is superior, followed by the CPM, finally by the RM in terms of errors on the validation data. Therefore, it is reasonable to conclude that the NN model is the most suited model for estimating the target cost. Further analysis of pertaining to NN models is discussed below.

#### 4.5 Further Analysis of NN Models

As mentioned in the previous section, the number of neurons set in the hidden layer was 5. It would be interesting to see the effects of varying the number of neurons in the hidden layer. Tables 8-10 show the effect for varying the neurons for Trials 1, 2, and 3 respectively.

**Table 8**

Maximum and average errors for Trial 1

		3	4	5	6	7	8	9	10
Training	Max Error	4.51	5.27	2.88	0.35	0.18	1.43	1.20	1.32
	Average Error	1.65	1.65	0.87	0.15	0.08	0.44	0.41	0.41
Validation	Max Error	9.15	8.79	8.78	9.87	11.04	9.05	9.10	9.31
	Average Error	6.39	6.30	5.55	6.99	7.87	6.32	6.33	6.45
Overall Average Error		2.75	2.72	1.65	1.73	1.87	1.80	1.77	1.80



**Table 9**

Maximum and average errors for Trial 2

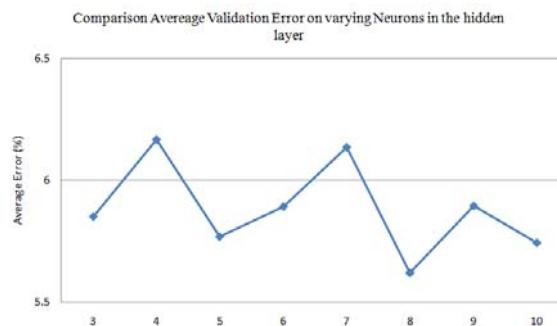
		3	4	5	6	7	8	9	10
Training	Max Error	7.07	5.86	4.91	0.00	18.46	18.79	19.40	2.85
	Average Error	1.53	1.48	1.21	0.00	9.76	9.54	10.15	1.11
Validation	Max Error	16.01	13.77	10.77	14.79	17.16	17.03	17.31	10.08
	Average Error	7.33	8.89	8.11	8.59	8.46	8.46	8.49	8.08
Overall Average Error		2.87	3.19	2.80	1.98	9.46	9.29	9.76	2.72

**Table 10**

Maximum and average errors for Trial 3

		3	4	5	6	7	8	9	10
Training	Max Error	7.24	7.30	0.00	2.15	3.58	0.93	4.26	5.73
	Average Error	2.18	2.01	0.00	0.47	0.78	0.20	0.94	1.23
Validation	Max Error	4.07	3.40	4.37	2.49	2.49	2.40	2.97	2.74
	Average Error	3.83	3.31	3.64	2.10	2.08	2.07	2.86	2.70
Overall Average Error		2.56	2.31	0.84	0.85	1.08	0.63	1.39	1.57

As can be seen from the sensitivity analysis, the number of neurons to select based upon the validation errors for Trials 1-3 would be 8, 10, and 5, respectively. In order to conduct the comparative analysis, one value is to be used for the number of neurons in the hidden layer. The value is set to 5 to minimize the number of weights created, to avoid over training. Moreover, as can be seen from Fig. 3, for different values for N, the error ranges between from 5.6% to 6.2%.

**Fig. 3.** Average error for combining Trails 1, 2, and 3

Thus, based upon this analysis it can be concluded for this case that there is no great sensitivity in varying the number of neurons in the hidden layer, contrary the claims of Huang and Huang (1991). Therefore, any selected value of N would yield reasonable results for the NN.

## 5. Conclusions and Future Research

When purchasing a new product, understanding the target cost is essential, as it will help in planning, bidding and make or buy decisions. The study presented in this paper is a case study to estimate the target costs of the main landing gears at Bombardier Aerospace. Even though this empirical study is for a specific application, the methodology utilized can be applied to various industries interested in estimating costs, as the model has the flexibility of being generalized.

The methods utilized to make the cost estimations are regression based parametric models and neural networks. This study shows that the overall performances of the neural network models are superior to that of the parametric models. In fact, the overall average error of all the trials presented in this paper of the regressions based parametric model is more than double that of neural networks.

Furthermore, the number of neurons in the hidden layer does not have a great impact on the performance of the model.

Even though this study has promising results, the analysis is based upon scarce data and is not an exhaustive study of all the possible sub-samples that can be made. If all the sub-samples are analyzed and the neural networks are computed for all of them, the average output obtained from all of the neural networks can be utilized as estimation for the target cost. Furthermore, the weights generated in the neural networks are based on the gradient descent method which can be trapped in a local optimal solution of the error surface. Other meta-heuristics, such as the genetic algorithm or simulated annealing can be utilized to obtain the weights and compared to the method presented in this paper. Moreover, the research can be extended by comparing the estimation of NNs to other techniques; such as complex non-linear models and fuzzy clustering.

## References

- Assaf, S. A. & Al-Hejji, S. (2006). Causes of delay in large construction projects. *International Journal of Project Management*, 24 (4), 349-356.
- Colmer, G., Dunkley, M., Gray, K., Pugh, P., & Williamson, A. (1999). Estimating the cost of new technology products. *International Journal of Technology Management*, 17(7-8), 840-846.
- Bashir, H. A., & Thomson, V. (2004). Estimating design effort for GE hydro projects. *Computer & Industrial Engineering*, 46(2), 195-204.
- Boehm, B. W. (1981). *Software Engineering Economy*. New Jersey: Prentice-Hall Inc.
- Bounds, G. (1998). The last word on project management. *IIE Solutions*, 30(11), 41-43.
- Frimpong, Y. (2000). Project management in developing countries: causes of delay and cost overruns in construction of groundwater projects. *Masters Research Project: University of Technology*, Sydney, Australia.
- Huang, S. C., & Huang, Y.-F. (1991). Bounds on the numbers of neurons in multilayer perceptrons. *IEEE Transaction on Neural Networks*, 2: 47 – 55.
- Kecman, V. (2001). *Learning and Soft Computing*. Cambridge: MIT Press.
- Kutner, M.H., Nachtsheim, C. J., Neter, J., & Li, W. (2004). *Applied Linear Statistical Models* (5<sup>th</sup> Edition), New York: McGraw-Hill/Irwin.
- Muralidhar, K., Parsa, R., & Sarathy, R. (1999). A general additive data perturbation method for database security. *Management Science*, 45(10), 1399-1415.
- Murmann, P. A. (1994). Expected development time reductions in the German mechanical engineering industry. *Journal of Product Innovation Management*, 11 (3), 236-252.
- Norris, K.P. (1971). The Accuracy of project cost and duration estimates in industrial R&D. *R&D Management*, 2 (1), 25-36.
- Pandya, A. S. & Macy, R. B. (1996). *Pattern Recognition with Neural Networks in C++*. Florida: CRC Press.
- Salam, A., Bhuiyan, N., Gouw, G.J., & Raza, S. A. (2009). Estimating design effort for the compressor design department: A case study at Pratt & Whitney Canada. *Design Studies*, 30(3), 303-319.
- Salam, A.M., Bhuiyan, N.F., Gouw, G.J., & Raza, S. A. (2008). A case study to estimate design effort for Pratt & Whitney Canada. *International Journal of Management Science and Engineering Management*, 3 (4), 284-293.
- Salam, A., Bhuiyan, N., Gouw, G., & Raza, A. (2008). *Measuring Design Effort in Product Development*. Germany: VDM Verlag.
- Snyman, J. A. (2005). *Practical mathematical optimization: an introduction to basic optimization theory and classical and new gradient-based algorithms*. New York: Springer.
- Walston, C. E., & Felix, C. P. (1977). A method of programming measurement and estimation. *IBM System Journal*, 16(1), 54-73.