Data Modelling and Optimization in Cost Estimation for Innovative Low Volume Product Development

Wai M. Cheung, Linda B. Newnes, Antony R. Mileham, Robert Marsh and John D. Lanham

Abstract—**This paper reports on the progress of the research and development of a data modelling and optimization method to support cost estimation in the product development process. This paper forms part of an investigation into Through-Life Costing of innovative low volume long life defence electronic systems. The paper briefly covers the literature review in the area of cost estimation in product development, in particularly the data sets needed to perform cost estimation and the method of modelling the data and the optimization techniques. The propose approach will be used to support cost estimation in product development decisions of innovative low volume product development.**

I. Introduction

FOR accurate and reliable cost estimates require historical data and information to be made. This is applicable in mass production and high volume products. As pointed out by Bode [1], cost estimation at the early stage of product development has always been difficult due to the availability of limited attributes. In addition, accurate and reliable cost estimation can only be obtained at a later stage of the development process when more information and data are presented. This is due to the fact that the cost models and systems used required a large amount of detailed data before a cost calculation could be made [2].

It is common knowledge that in each phase of the development of a product, a company spends money and also incurs costs. The Ford Motor Co Ltd estimates that although the design stage only constitutes 5% of the total product cost, the influence on the total product development cost is as high as 70% [3]. Therefore, the final cost of a product under development is usually an important design attribute as illustrated in Figure 1. It is thus essential to understand the value of this cost design attribute as early as possible in the design cycle and preferably in the conceptual design stage. The more the project is advanced the greater the difficulty of reducing the final cost because of the high costs of modification and change [4].

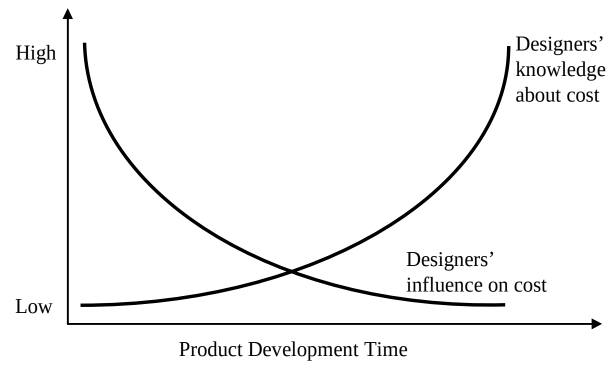

Fig. 1. Cost Estimation Dilemma in Product Development [1]

Over the years many researchers have made considerable research efforts to tackle the problems of cost estimation at the earliest stage of product development. Some typical approaches and techniques used to support decision making in cost estimation of product development are parametric, product family, variant, analytical, learning curve, neural network and life cycle costing [1, 5-11].

However, these methods in general would not be applicable for innovative low volume products. In comparison, cost estimation in low volume innovative product development is usually hindered by the lack of statistically significant data and reliable methods. This paper therefore proposes a new approach of performing data modelling and optimization to address cost estimation in innovative low volume product development. The layout of this paper is as follows: Section 2 presents the literature review of data modelling and optimization; Section 3 describes the data needed to support the new approach; Section 4 discusses the data sets and searching mechanism; Section 5 discusses the methods of data optimization and finally the conclusion and further work.

II. Literature Review

Research that addresses cost estimation at the early stage of product development using traditional cost estimation and optimization techniques are as follows.

Bode [1] compared the performance of neural networks and other conventional cost estimation methods at the early stage of bearing development. Neural networks are non-parametric techniques have the ability to be trained and learned to perform accurate estimation with limited attributes that fit non-linear curves. However, the technique

Manuscript received March 28, 2009. This work is supported and funded by the Engineering and Physical Sciences Research Council (EPSRC), the Innovative electronics Manufacturing Research Centre (IeMRC) (GR/T07459/01) and the University of Bath's Innovative design and Manufacturing Research Centre (IdMRC) (EP/E00184X/1).

W.M. Cheung is with the IdMRC, University of Bath, Department of Mechanical Engineering, Bath, UK. (Phone: +44 (0)1225 386131; e-mail: w.m.cheung@bath.ac.uk).

L.B. Newnes is with the IdMRC, University of Bath, Department of Mechanical Engineering, Bath, UK. (e-mail: l.b.newnes@bath.ac.uk).

A.R. Mileham is with the IdMRC, University of Bath, Department of Mechanical Engineering, Bath, UK. (e-mail: a.r.mileham@bath.ac.uk).

R. Marsh is with Bristol Institute of Technology, The University of the West of England, Bristol, UK. (email: robert.marsh@uwe.ac.uk).

J.D. Lanham is with Bristol Institute of Technology, The University of the West of England, Bristol, UK. (email: john.lanham@uwe.ac.uk).

relies on past case data. Tu et al [7] presented a cost data index structure with two traditional cost estimation methods namely, generative and variant cost estimation, for the development of a computer-aided cost estimate and control system in mass customization of sheet metal products. For cost optimization, an optimal algorithm for the selection of alternative operation routines and suppliers was developed using the dynamic programming technique. However the methodology relies on past knowledge and experiences.

Kaufmanna et al [12] proposed an optimization framework to minimize the direct operating cost at a part level based on the cost/weight optimization of composite aircraft structures. The framework has been implemented based upon the parametric method. Shehab and Abdalla [13] developed a prototype object-oriented and rule-based system for product cost modelling and design for automation to support decision making at the early design stage. The prototype system was implemented in a combination of heuristics data, algorithmic approach and fuzzy logic techniques. The system allows users to generate accurate cost estimates for new designs by exploring alternative materials and production processes. Deiab and Al-Ansary [14] developed a systematic multi-phase procedure to optimize the design and manufacturing parameters using the genetic algorithm method. The aim was to minimize the total manufacturing cost under dimensional, weight, and machine power constraints. Similar to parametric methods, the approximations are based on past case data where cost is known.

The research projects mentioned in this review are relevant to high volume products, where typically, past and historical data are available to support their methodologies. The next section of this paper will focus on the discussion of the data needed in product development of innovative low-volume long-life electronic products.

III. Type of Data for Cost Estimation in Low Volume Innovative Product

The focus of this paper is to propose a method of utilizing data and rule-based techniques to optimize the performance of cost estimation of defence electronic systems at the early stages of a product development process. A modular approach of constructing a Bill-Of-Materials (BoM) of a product from a 'Digital Library' has been used. The Digital Library is a data structure used to capture cost data and information of a set of domains namely: mechanical, electronics modules/components, product, process and resource [15]. The BoM of the new product is made up by a set of elements such as cost models and cost data from the Digital Library. The decision on what elements the model will be constructed from is based on a number of questions, and the answers to these questions may be obtained by implementing, for example, a rule-based "production system". To establish what rules should be used to implement such a system, two further primary questions need to be met:

a) Is there an existing cost model or data in the library for a new product?

b) How relevant is the data in the library? (The degree of relevance will be used to determine the accuracy of the estimation.)

Only after the above analysis, will the elements from the library be used or be acceptable in a new product configuration. If the elements were not be applicable, it should be established what other alternatives could be used. Thus, this could be solved either:

a) Construct the model from finer detail (e.g. sub-modules of a BoM), or

b) Choose an element that is the nearest match.

This fundamental aspect is the basis of deriving the data sets that are used to support the data searching method in the cost estimation process.

IV. The Data Sets to Support the Data Searching Mechanism

A study has been carried out on how industry utilizes data to perform cost estimation. There are five types of data set that have been identified which could influence the impact of the accuracy and confidence levels of the cost estimated. The data set is refereed to as the Data Searching Mechanism (DSM) as represented in Figure 2.

The rules that are used in the DSM are built upon the data sets which are explained as follows:

1) Commercial-Off-The-Shelf (COTS) - fixed standard cost data from supplier.

2) Parametric - cost outputs that are directly dependent upon the characteristics or parameters of the product such as weight, volume, length and the number of inputs/outputs. In the case of electronic products, for example, 'material type', 'size of PCBs', 'number of components (resistors, ICs) etc

3) Variant - in general, new variants of current products usually involve incremental changes rather than a novel/new design. Thus, variant design applies a 'product family' approach. Cost is derived from a mixture of existing and old products.

4) Detailed - when parametric data is unavailable, or a more accurate result is required and a number of sub-systems exist for further cost analysis, cost data can be obtained using a Bill-of-Materials (BoM) approach in which the available BoM for the selected design enables historical data to be accessed.

5) If new and uncertain technology is involved then the cost can be modelled using:

(a) Monte Carlo simulation to analyse the uncertainty. This analysis is based on input values from a range of similar products in order to simulate a probability distribution. Thus 'New Technology against Cost' characteristic can be obtained.

(b) The application of 'Monte Carlo' simulation depends on similar technologies and historical data. If there is no information or data available to support 'Monte Carlo' simulations, another approach is needed. As illustrated in Figure 2, this research has adapted the 'technology

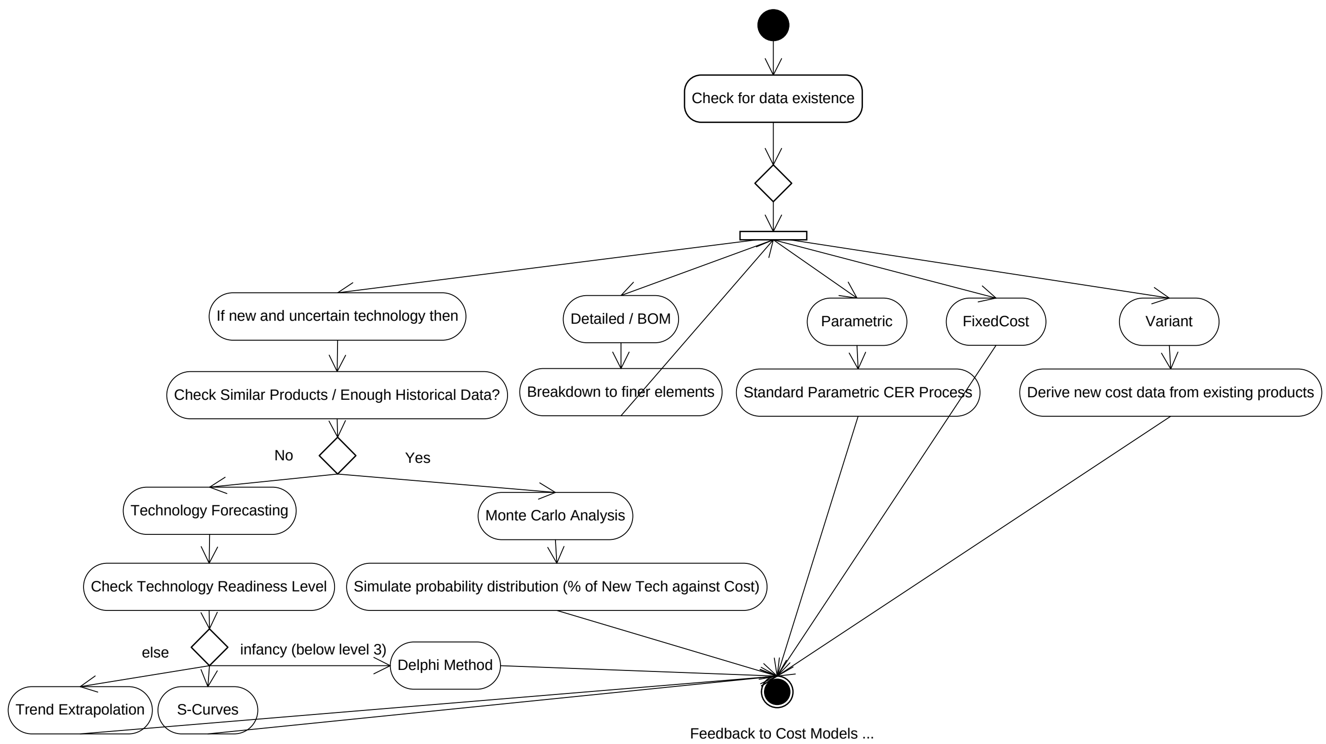

Fig. 2. Rules for Data Searching Mechanism

forecasting' method to evaluate the impact of this scenario. The Technology forecasting technique depends on 'technology readiness level' (TRL) [16]. According to Moorhouse [17], TRL consists of *nine* levels from 'basic technology research' to the technology ready to be 'launched and operated'. In order to check the impact of emerging technologies, there are three techniques available:

(1) If the emerging technology is in its infancy stage, i.e. below level 3, the application of the 'Delphi Method' should be used to predict the growth pattern of new and uncertain technologies, thus, by 'quantifying' the prediction, the data can be used to support a cost model.

(2) If the emerging technology is above level 3, techniques such as 'S-Curves' and 'Trend Extrapolation' will be used. S-Curves are a mathematical modelling technique used for analysing technological cycles and predicting the introduction, adoption and maturation of innovations [18]. Trend extrapolation uses 'correlation studies' and 'analytical models' to see what will happen in the near future [16].

The method and techniques used to implement the DSM are discussed in the following section.

V. The Data Modelling and Optimization Approach

A. *Data Algorithm for Optimization*

Figure 3 illustrates an example of a modular approach in cost estimating of an electronics system. The figure indicates that a system has a number of subsystems, and a subsystem could consist of different kinds of component and assembly operations. Therefore, a subsystem could consist of different types of cost data. It is important that an intelligent search method is used to optimize the availability of the cost data so that the cost can be accurately predicted.

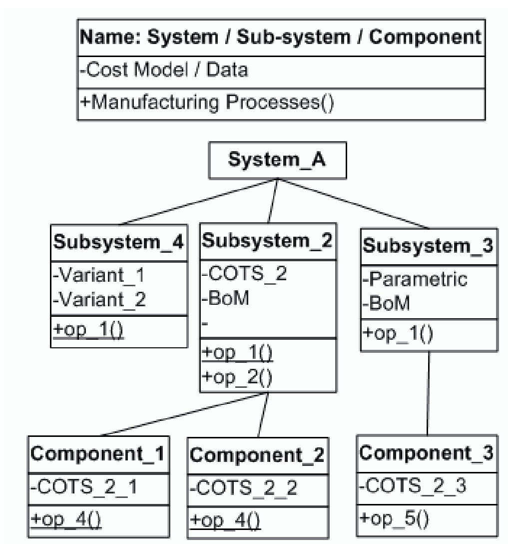

Fig. 3. Example Modular Approach

The definition of the 'data searching algorithm' is shown in Figure 4 which is represented as 'if/then' statements. The algorithm represents the procedural steps of performing data searching, and these procedural steps are based on the set of rules in the DSM. For example, in order to search what kind of data is used in 'subsystem_2' (as shown Figure 3), a search will begin to check the existing data type as follows:

1). If it is 'variant' then return a new value based on 'existing product data', this could help the designer to adapt or modify the nearest existing product to perform cost estimation.

2). If it is 'COTS' then return a new value based on a 'fixed price'.

3). If it is 'Parametric', derive the cost from a Cost Estimation Relationship (CER).

4). If it is a 'detailed/BoM' then the search will reiterate again or perform a detailed design to accurately predict the cost in the early stages.

5). If it is an 'uncertain technology' then 'run a Monte Carlo simulation' or run the 'technology forecasting technique'.

```
Run through the conditionals of the datasets
  then
    beginSearch
      x = 1                                                  Declarations
      value
      NewCostData
        {
        if Concept&Subsystem(x)&(Variant) exists then
          Variant = True                                     Checking
          value(x) = NewCostData(x)                          for
           NewCostData(x) = (ExistingProduct(x) +            variant
                             ExistingProduct (x+1) + …
                             ExistingProduct (n)) / n
          return value
        else
          if Concept&Subsystem(x)&(COTS)exists then          Checking
           COTS = True                                       for
           value(x) = NewCostData(x)                         fixedCost
          return value
        else
          if Concept&Subsystem(x)&(Parametric) exists
                then                                         Checking
           Parametric = True                                 for
           value(x) = NewCostData(x)                         parametric
           NewCostData(x) = Derived from Cost Estimation
                                  Relationship (CER)
          return value
        else
          if Concept&Subsystem(x)&(Detailed/BOM) exists      Checking
                then                                         for
            Detailed/BOM = True                              Detailed/
          {                                                  BOM
            value(x) = NewCostData(x)
            breakdown to finer elements                   Checking for
                begin search on                           variant,
                  'variant', 'parametric', 'fixedCost',   parametric,
                      'Uncertain tech', 'BoM'.            fixedCost and
                end search                                UncertainTechs
           }                                              in the existing
          return value                                    BoM
        else
          if Concept&Subsystem(x)&(UncertainTech) exists
                then
            UncertainTech = True
          {
            Check for historical data                        Checking
                If 'true' then                               for
                      {run Monte Carlo Simulation}           Uncertain
                else                                         Tech
                      {technology forecasting techniques}
          }
         check the statistical significance of the result
      end if
    x = x+1
  end search
end run
```

Fig. 4. The Rule-Based Searching Algorithm

The implementation of the DSM is under development and the techniques used for its development is discussed in the following section.

B. Implementation of the DSM

The algorithm of the DSM has been developed using rule-based techniques [19]. Currently, the DSM has been implemented using a forward chaining method (also known as data-driven reasoning) using a First In, First Out queue processing technique [20]. As illustrated in Figure 5, the purpose of the DSM is used to search for relevant cost information during the early design configuration process. It is at this stage that the cost of alternative design concepts needs to be determined and as precisely as possible to support a designer's decision making in the product development process.

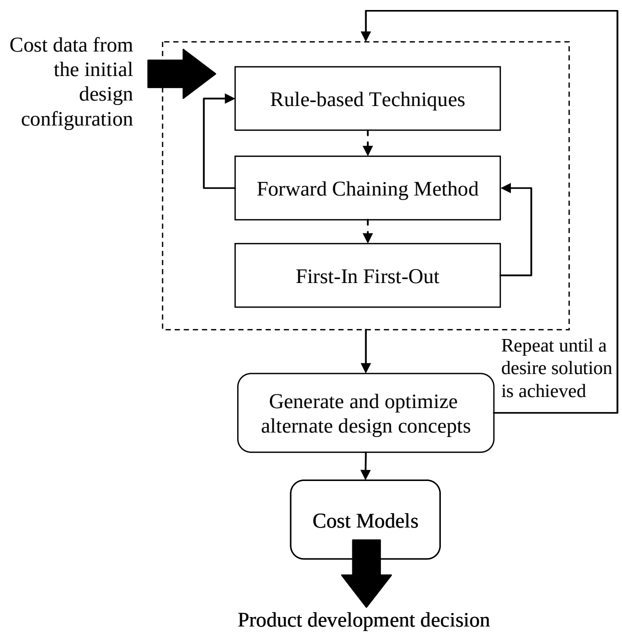

Fig. 5. Optimization Method

A pilot demonstrator has been implemented as shown in Figure 6, where rule evaluation is fired from the button on the Graphical User Interface. This utilizes standard inference engine with if/then statements. The inference engine then matches the agenda contents to the rules in the rule-base(s). If they match, the appropriate rules are fired which produce new facts on the agenda. This mechanism repeats until the agenda is blank or the goal is satisfied. A goal is used in the case of a known model being selected. The percentage (%) new product is used to select the best kind of model. If some of the product is new then a variant approach is favoured. The novelty of the approach is a combination of (1) a data driven, (2) multi-model selection process, and (3) an evaluation of the models in parallel with a measure of relevance to the available data attached.

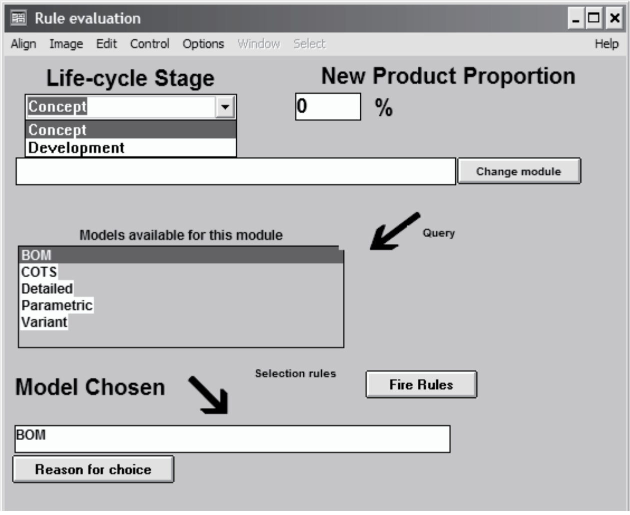

Fig. 5. Optimization Method

VI. Conclusions and Future Work

A new approach in data modelling and optimization to support cost estimation for low volume innovative products has been proposed. The method is currently under development and will be applied to address product development in defence electronic systems. The new and uncertain technologies evaluation has also been defined which will be implemented as part of the DSM. Further research to explore is (1) the utilization of the result from S-Curves and Trend Extrapolation to feed into the cost models and (2) to quantify the expert opinions from the Delphi method into numerical form so that the result can be used to support cost estimation.

Acknowledgment

The authors would like to thank the following for assisting their research activities in this area: the industrial collaborators, MOD DE&S Supplier Engagement Team - Terry Johns and GE-Aviation.

References

[1] J. Bode, "Neural networks for cost estimation: simulations and pilot application", International Journal of Production Research, 2000, vol.38, no. 6, pp. 1231-1254.

[2] W.M. Cheung, R. Marsh, L.B. Newnes, A.R. Mileham, and J.D. Lanham, "Standards and inference of design information in through-life costing of defence electronic products", in the Journal Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, Publisher: Professional Engineering Publishing, 2009, vol. 223, no. 2, pp. 169 – 181, ISSN 0954-4054, DOI 10.1243/09544054JEM1315.

[3] S.M. Fohn, A. Greef, R.E. Young, and P. O'Grady, *Concurrent Engineering*, Book Series: Information Management in Computer Integrated Manufacturing, Publisher Springer Berlin / Heidelberg, 1995, Vol. 973, pp. 493-505, ISBN 978-3-540-60286-6, DOI 10.1007/3-540-60286-0_116.

[4] D.G. Ullman, *The Mechanical Design Process*, Publisher: McGraw-Hill Higher Education; 3rd Edition, 416 pages, ISBN-10: 0071122818 ISBN-13: 978-0071122818.

[5] A.R. Mileham, G.C. Currie, A.W. Miles and D.T. Bradford, "A parametric approach to cost estimating at the conceptual stage of design, Journal of Engineering Design, 1993, vol. 4, no. 2, pp.117-125.

[6] B. Verlinden, J.R. Duflou, P. Collin and D. Cattrysse, "Cost estimation for sheet metal parts using multiple regression and artificial neural networks: A case study", International Journal of Production Economics, 2008, vol. 111, pp. 484–492.

[7] Y.L. Tu, S.Q. Xie and R.Y.K Fung, "Product development cost estimation in mass customization", IEEE Transactions on Engineering Management, 2007, vol. 54, no. 1, pp. 29-40.

[8] J. Jiao and M. M. Tseng, "Understanding product family for mass customization by developing commonality indices", Journal of Engineering Design, 2000, vol. 11, no. 3, pp. 225–243.

[9] T.L. Smunt, "Log-linear and non-log-linear learning curve models for production research and cost estimation," International Journal of Production Research, 1999, vol. 37, no. 17, pp. 3901–3911.

[10] K. Tamas, S. Lozano, F. Guerrero, and L. Onieva, "A flexible costing system for flexible manufacturing systems using activity based costing", International Journal of Production Research, 2000, vol. 38, no. 7, pp. 1615–1630.

[11] J.H. Park and K.K. Seo, "Incorporating life-cycle cost into early product development", in the Journal Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, Publisher: Professional Engineering Publishing, 2004, vol. 218, no. B9, pp. 1059–1066.

[12] M. Kaufmanna, D. Zenkerta and C. Matteib, "Cost optimization of composite aircraft structures including variable laminate qualities", Composites Science and Technology, 2008, vol. 68, no. 13, pp. 2748-2754.

[13] E. M. Shehab and H. S. Abdalla, "A design to cost system for innovative product development", in the Journal Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture , Publisher Professional Engineering Publishing, 2002, vol. 216, no. 7, pp. 999-1019, ISSN 0954-4054,DOI 10.1243/09544050260174201.

[14] I. Deiab and M. Al-Ansary, "Concurrent optimization of manufacturing cycle cost by genetic algorithms", in the Journal Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 2007, vol. 221, no. 5, pp. 821-834, ISSN 0954-4054, DOI: 10.1243/09544054JEM752.

[15] W.M. Cheung, L.B. Newnes, A.R. Mileham, R. Marsh and J.D. Lanham, "Through life costing in defence electronic systems: an integrated data-driven multi-level approach", in the Proceedings of the 38th International Conference in Computer and Industry Applications, CIE2008, 31st October – 2nd November, 2008, Beijing, China, vols. 1-3, pp. 2119-2128.

[16] M.R. Kirby and D.N. Marvis', "Forecasting technology uncertainty in preliminary aircraft design' SAE 1999 Transactions, Journal of Aerospace, 1999, vol. 108, pp. 1388-1399.

[17] D.J. Moorhouse, "Detailed definitions and guidance for application of technology readiness levels", Journal of Aircraft, 2002, vol. 39, no. 1, pp. 190-192.

[18] M.S. Slocum and C.O. Lundberg, "Technology forecasting: from emotional to empirical", Creativity and Innovation Management, Blackwell Publishers Ltd, 2001, Vol. 10, No. 2, pp. 139-152, , DOI: 10.1111/1467-8691.00215.

[19] E. Turban, J.E. Aronson and T.P. Liang, *Decision Support Systems and Intelligent Systems*, Publisher: Prentice-Hall, Inc, 7th Edition, 2004, ISBN: 0130461067.

[20] J. Finlay and A. Dix, *An Introduction to Artificial Intelligence*, UCL Press, Taylor & Francis Group, 1996, Pages 276, ISBN: 1-85728-399-6.