Data Mining QFD for The Dynamic Forecasting of Life Cycle under Green Supply Chain

CHIH-HUNG HSU[1*], AN-YUAN CHANG[2], HUI-MING KUO[3],
[1*]Institute of Lean Production Management, Hsiuping Institute of Technology
11, Gungye Road, Dali, Taichung County 412, TAIWAN, R.O.C.
[*]Corresponding author's e-mail: chhsu@mail.hit.edu.tw http://chhsu.tw

[2]Institute of Industrial Engineering and Management, National Formosa University
No.64, Wunhua Road, Huwei, Yunlin County 632, TAIWAN, R.O.C.

[3]Department of Logistics Management, Shu-Te University
59, Hun Shan Road, Yen Chau, Kaohsiung County 824, TAIWAN, R.O.C.
ayc@nfu.edu.tw[2] hmkuo@mail.stu.edu.tw[3]

Abstract: - The satisfaction of customer requirements is critical issue for the computer designers and manufacturers, because computer design is a high risk and value-added technology. When considering green design, designers should incorporate the voices from the customers and because they are the driving force. On the other hand, data mining from large marketing database has been successfully applied in a number of advanced fields. However, little study has been done in the quality function deployment of identifying future customer requirements for computer design and manufacture, using data mining. This study uses data mining cycle in QFD to forecast future customer requirements for green design of life cycle. The use of time series-based data mining cycle to predict the weights is advantageous because it can (1) find the future trend of customer requirements; (2) provide the computer designers and manufacturers with reference points to satisfy customer requirements in advance. The results of this study can provide an effective procedure of identifying the trends of customer requirements and enhance dynamic forecasting of life cycle under green supply chain in the computer marketplace.

Key-Words: - Data mining, Quality function deployment, Customer requirements, Dynamic forecasting, Life cycle, Green supply chain.

1 Introduction

Green products - products that can reduce the load on the environment during use and disposal – have additional marketing entreaty. Recycling of materials, and adequate reuse of subassemblies can greatly reduce waste generation, thus increasing product green compatibility. Previous researches have been reported in recycling analysis, disassembly sequences analysis, and disassembling planning [8, 25, 28]. Also, the most approved way to evaluate the environmental profile of a green product is the Life Cycle Assessment (LCA) methodology [9]. According to the LCA evaluation, several green products are evaluated and produced. However, most green products are not valuable in the market as expected since those products are focused entirely on environmental impact analysis without regarding to the customer need. Therefore, the customers might be unwilling to buy those products even though they understand buying the green products could reduce the environmental impacts [29].

Changing design procedures is particularly difficult because product designers may face many uncertainties, conflicting objectives, and cost effectives. Rarely in one product greener in every dimension than other products; there are usually tradeoffs among features [28, 26, 31]. Therefore, when considering green design, designers should incorporate the voices from the customers and because they are the driving force.

On the other hand, the application domain of data mining is quite broad and plausible in health insurance [30], surface roughness prediction [4], biomedical technology [24], risk prediction [16], human resource management [13], semiconductor manufacturing [3], production schedule [7], marketing [17], domain ontologies [20], customer lifetime value [5] and others. However, little research has also been applied to identify customer requirements (CRs) and weights in QFD using data mining. This study applied a time series-based data mining cycle for dynamic forecasting of product life cycle, using sales questionnaire database, to identify future CRs for the computer designers and manufacturers. By applying the proposed approach, future CRs can be found from a large database to enhance their competitiveness in the computer marketplace. This procedure help manufacturers reduce the use of material, work force, money, energy, etc., and the greatest impact is to drastically shorten the cycle time of product production. Finally, customer opinions are actually listened and responded by the best alternative in the drawing board such that green compatibility can be considered in the early phases.

2 Related Research

2.1 Quality function deployment

Quality Function Deployment (QFD) is a Japanese development and design technology [15]. QFD was first introduced by Akao in 1972 at Mitsubishi's Kobe shipyard site, and then Toyota and its suppliers developed it further for a rust prevention study [10]. After the concept of QFD was introduced in the US through auto manufacturers and parts suppliers [18], many US firms, such as AT&T, Digital Equipment, Ford, GM, Hewlett-Packard, Procter & Gamble, and Raychem, applied QFD to improving communication, product development [1, 2].

QFD has been widely applied to achieve CRs and improve customer satisfaction in many fields. Some researchers defined QFD as follows: "This technology focuses and coordinates skills within an organization, first to design, then to manufacture and market products that customers want to purchase and will continue to purchase [6]." Some companies have claimed great success with QFD. Proponents assert that QFD has helped them reduce production costs, design time and cost; increase customer satisfaction and product quality [2, 27].

QFD is a cross-functional planning tool which is used to ensure that the voice of the customer is deployed throughout the product planning and design stages. QFD is used to encourage breakthrough thinking of new concepts and technology. Its use facilitates the process of concurrent engineering and encourages teamwork to work towards a common goal of ensuring customer satisfaction. Because the voice of the customer is essential, the House of Quality (HOQ) converts each CR into one or more engineering characteristics (ECs) in the first phase of QFD. The main goal of HOQ is to identify CRs and weights for the product (WHATs) and then to convert these needs into ECs (HOWs). The components of HOQ are shown in Figure 1.

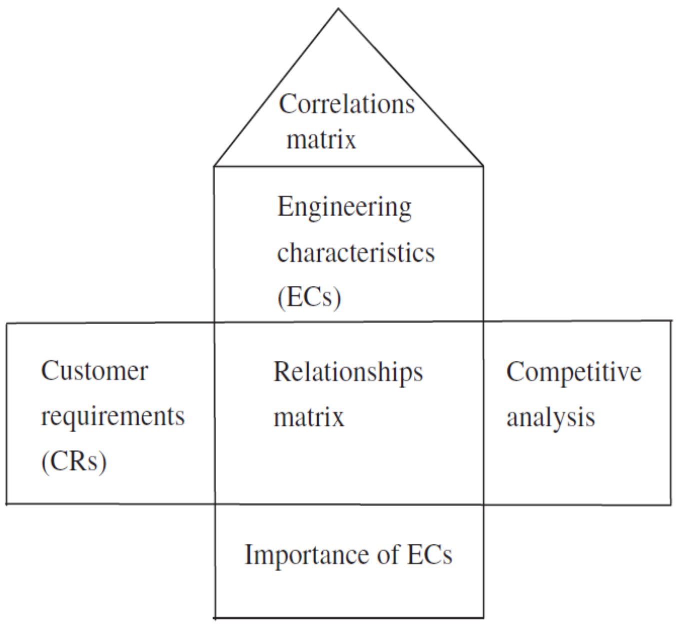

Fig. 1 Components of HOQ.

2.2 Life Cycle Assessment

Zust and Wagner [24] explained four phases of the product life cycle: (1) product definition, (2) product development, (3) product manufacturing and marketing, and (4) product usage. At each of these phases, there exists a definition of objectives, activities, and deliverables for the next phase. Keys [16] explained that during the conceptual model phase, various design models of the product are generated. From these conceptual models, requirements, and specifications will evolve decisions for breadboard and brassbound models. Within the LCA framework, the greatest challenge is the assessment of the impacts associated with environmental releases during the manufacturing, transportation, use and disposal of products. Also, the Hewlett-Packard addressed the life-cycle issue by prototyping software, defining development and phases, and standardizing modules and packages [29].

3 Data Mining and Time Series

Berry and Linoff defined data mining as the analysis of huge amounts of data by automatic or semi-automatic means, in order to identify significant patterns or rules [6, 19]. One of the most important data mining techniques is time series analysis. Time series data often arise when monitoring industrial processes or tracking corporate business metrics [22].

Time series analysis can be used to accomplish different goals:

(1). Descriptive analysis determines what trends and patterns a time series has by plotting or using more complex techniques.

(2). Spectral analysis is carried out to describe how variation in a time series may be accounted for by cyclic components. This may also be referred to as "Frequency Domain". With this an estimate of the spectrum over a range of frequencies can be obtained and periodic components in a noisy environment can be separated out [23].

(3). Forecasting can do just that - if a time series has behaved a certain way in the past, the future behavior can be predicted within certain confidence limits by building models.

(4). Intervention analysis can explain if there is a certain event that occurs that changes a time series. This technique is used a lot of the time in planned experimental analysis.

(5). Explanative analysis using one or more variable time series, a mechanism that results in a dependent time series can be estimated [11].

One of the most important forecasting techniques is exponential smoothing analysis for time series analysis. Forecasts generated with this method are a weighted average of the past values of the variable. The weights decline for older observations. The rationale is that more recent observations are more inuential than older observations. The exponential smoothing analysis is,

$$F_t = F_{t-1} + \alpha\left(A_{t-1} - F_{t-1}\right) \tag{1}$$

F_t : *The forecast value for period t*

F_{t-1} : *The forecast value for period t-1*

A_{t-1} : *The actual value for period t-1*

α : *Smoothing coefficient* $0 \le \alpha \le 1$

The higher the value of alpha the more weight is given to current values [12].

Thus, this study proposed a time series-based data mining cycle, in order to mine the patterns of weights for identifying future CRs in QFD.

4 The Data Mining Procedure

This study uses data mining cycle to identify future CRs with each respective step closely involved. The data mining cycle involves a series of activities, from defining the problem to evaluating and applying the results. The previous steps can be served as the baseline reference for the next step, and the steps for dynamic forecasting of life cycle under green supply chain are described below.

4.1 Defining the problem for data mining

Owing to unknown weights for future CRs, a large marketing database was created by a professional computer manufacturer in Taiwan, based on many sales questionnaries, according to four period questionnaires; this resulted in a huge amount of data.

The intent of this study was to explore and analyze a huge amount of data, by employing a time series-based data mining cycle in QFD, so as to identify the weights within customer questionnaires in each period. Based on these the weights of CRs, the future CRs for dynamic forecasting of life cycle under green supply chain may be discovered and the results can be encouraged and beneficial the computer designers and manufacturers.

4.2 Data preparation and analysis

The data was processed, and analyzed, in order to enhance the efficiency and ensure the accuracy of the results [21]. Before mining the data, it had to be checked and processed, with all abnormal or missing data being separated out. As a result, of the 18,000 questionnaires, 363, which had missing or abnormal data, were deleted; this left a total of 17637 valid questionnaires regarding the operating CRs. There are night CRs and ten ECs for each questionnaire, as shown in Table 1 and Table 2. The green QFD matrix for computer is as shown in Figure 2.

Table 1 Definitions of customer needs

Voice of customer	Customer Requirements
CR1	Speed
CR2	Recyclable
CR3	Energy Saving
CR4	Easily Disassembly
CR5	Easily Maintenance
CR6	No Toxical Material Released
CR7	Size or Weight
CR8	Operating Quality
CR9	Price or Cost

Table 2 Definitions of design requirements

		Voice of Engineering	Engineering Characteristics
Life Cycle Phases	Raw Material	EC1	Material Reduction
		EC2	No Dangerous Material
	Manufacturing	EC3	Pollution Control
		EC4	Low Energy Exhausting
	Disassembly	EC5	Modularization
		EC6	Tools Usage
	Transportation	EC7	Package Reduction
	Usage	EC8	Energy Saving
		EC9	Maintenance
	Disposal	EC10	Reuse or Recycled

		Life Cycle Phases									
		Raw Material		Manufacturing		Disassembly		Transportation	Usage		Disposal
	Weights	EC1	EC2	EC3	EC4	EC5	EC6	EC7	EC8	EC9	EC10
CR1			3			9	3		9	1	1
CR2		9		3	9	9	1	9	3	9	9
CR3		3		1	9	9		3	9	9	1
CR4		1			3	9	3		1	3	3
CR5		3	1	1	3	3	9		3	9	9
CR6		1	9	9		1	1			9	3
CR7		3			1		1	3	9	3	3
CR8						1	3		3	3	
CR9		9			3	1	1	1	9	3	9
Importance of EC											

Fig. 2 The green QFD matrix for computer.

4.3 Data mining by time series method

The weights of four periods for each CR are periodically computed in Table 3. The weight for each CR is evaluated by a 1–10 scale, where a CR with a lower value is not more important. On the other hand, it is essential for the company to reflect CRs by corporate language and then fulfil those ECs to satisfy CRs. When CRs are translated by HOWs, the computer designers and manufacturers have to check the relationship between WHATs and HOWs.

QFD represent the respective strong (with a weight of 9), moderate (with a weight of 3), and weak relationship (with a weight of 1), while the blank is zero. Taking period 1 as an example, the matrix relationship between CRs and ECs is shown in Table 4.

Table 3 The weights of CRs

	Period 1	Period 2	Period 3	Period 4
CR1	6.3	5.4	5.7	6.8
CR2	5.6	5.3	6.4	7.2
CR3	7.2	6.2	7.2	7.4
CR4	4.8	5.8	5.9	6.1
CR5	5.4	5.6	6.5	7.5
CR6	6.7	5.1	5.2	6.8
CR7	6.3	5.4	7.6	6.4
CR8	5.9	6.7	6.4	6.2
CR9	6.2	5.3	5.1	6.5

Table 4 The HOQ of period 1

	Weights	EC1	EC2	EC3	EC4	EC5	EC6	EC7	EC8	EC9	EC10
CR1	6.3		3			9	3		9	1	1
CR2	5.6	9		3	9	9	1	9	3	9	9
CR3	7.2	3		1	9	9		3	9	9	1
CR4	4.8	1			3	9	3		1	3	3
CR5	5.4	3	1	1	3	3	9		3	9	9
CR6	6.7	1	9	9		1	1			9	3
CR7	6.3	3			1		1	3	9	3	3
CR8	5.9					1	3		3	3	
CR9	6.2	9			3	1	1	1	9	3	9
Importance of EC		174.4	84.6	89.7	170.7	250.1	124.4	97.1	289.5	300	221.7

Through checking the relationship between WHATs and HOWs, the matrix relationship between CRs and ECs were determined. Subsequently, data mining was undertaken, using a time series-based data mining cycle, to mine the weights and determine the trend of each CR for the next period.

According to the data mining cycle, the predicted weights of CRs in the next period (period 5) would be estimated as shown in Table 5. As shown, the weight of CR1 in the period 5 is 6.2; thus, these predicted weights for the CRs were chosen for the next stage of processing.

Table 5 Predicted weights for the CRs in the period 5

	Period 1	Period 2	Period 3	Period 4	Period 5 predicted weights
CR1	6.3	5.4	5.7	6.8	6.2
CR2	5.6	5.3	6.4	7.2	7.1
CR3	7.2	6.2	7.2	7.4	7.8
CR4	4.8	5.8	5.9	6.1	6.1
CR5	5.4	5.6	6.5	7.5	7.2
CR6	6.7	5.1	5.2	6.8	6.9
CR7	6.3	5.4	7.6	6.4	6.8
CR8	5.9	6.7	6.4	6.2	6.1
CR9	6.2	5.3	5.1	6.5	6.9

4.4 Evaluation and Application of Results

This study uses data mining cycle in QFD to forecast future CRs. In addition, the time series-based data mining cycle can be applied to predict the weights and determine the trend of each CR for the next period. The use of time series-based data mining cycle to predict the weights is advantageous because it can (1) find the future trend of CRs; (2) provide the computer designers and manufacturers with reference points to satisfy CRs in advance. The results of this study can provide an effective procedure of identifying the trends of CRs and enhance customer relationship management in the computer marketplace.

Owing to the entire forecast methods have forecast errors, the Mean Squared Error (MSE) and control charts of forecast were applied to monitor the accuracy of forecast in this study. The MSE is the average of the squared forecast errors. Forecast error is defined as the difference between actual value and the forecast.

$$MSE = \left[\sum_{i=1}^{n} (A_i - F_i)^2 \Big/ n-1 \right] \quad (2)$$

Taking the CR1 as an example, the mean squared error is 0.68. Furthermore, the study calculated the double control limit for the control charts of forecast. Because the upper control limit is 1.65 and lower control limit is -1.65 in the study, all of the forecast errors of CR1 were under the double control limit, as shown in Table 6. The forecast errors in other CRs are less than the double control limit. Thus, the exponential smoothing analysis is clearly quite accurate.

Table 6 The forecast error for CR1

	Period 1	Period 2	Period 3	Period 4	Period 5 predicted weights
At	6.3	5.4	5.7	6.8	6.7
Ft	N/A	6.3	6.1	5.9	6.2
Forecast error = At - Ft	N/A	-0.9	-0.4	0.9	0.5
$\lvert e \rvert$	N/A	0.9	0.4	0.9	0.5
e^2	N/A	0.81	0.16	0.81	0.25

To gain a better insight into the predicted weights among the nine CRs resulting from the time series-based data mining cycle, a line plot was drawn of the weights of the CRs. As can be seen in Figure 1 and Figure 2, the weights bear marked differences in the CRs for each period. The trend for each CR can be understood and controlled by the computer designers and manufacturers with the information in Figure 1 and Figure 2. The importance of each CR can be considered to know future CRs. The computer designers and manufacturers can be designed and planned to satisfy with customer future CRs in advance.

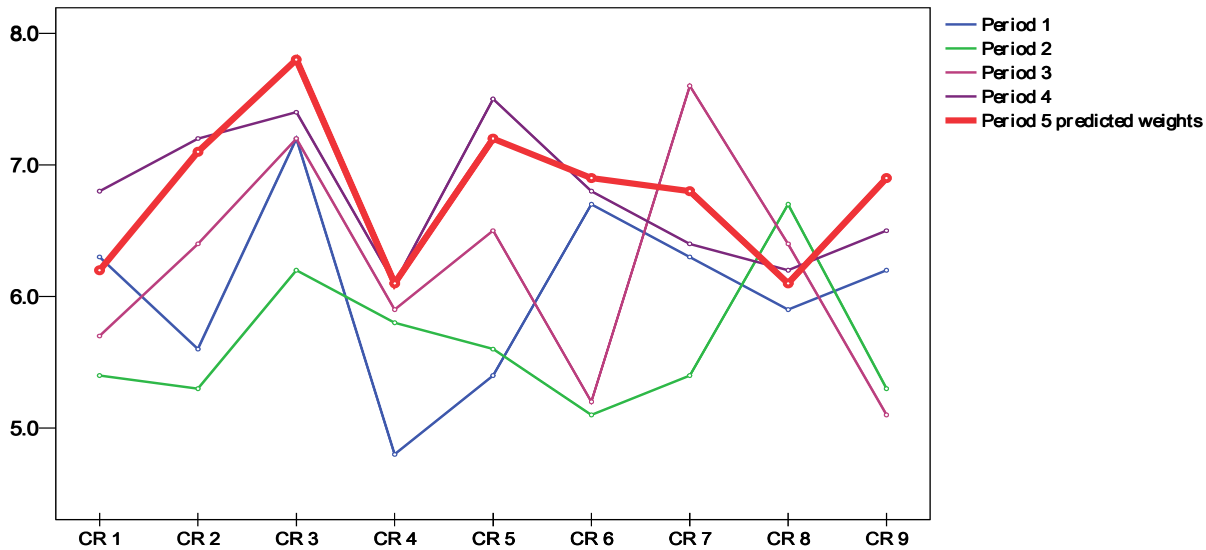

Fig. 1 The weight plot of CRs for each period.

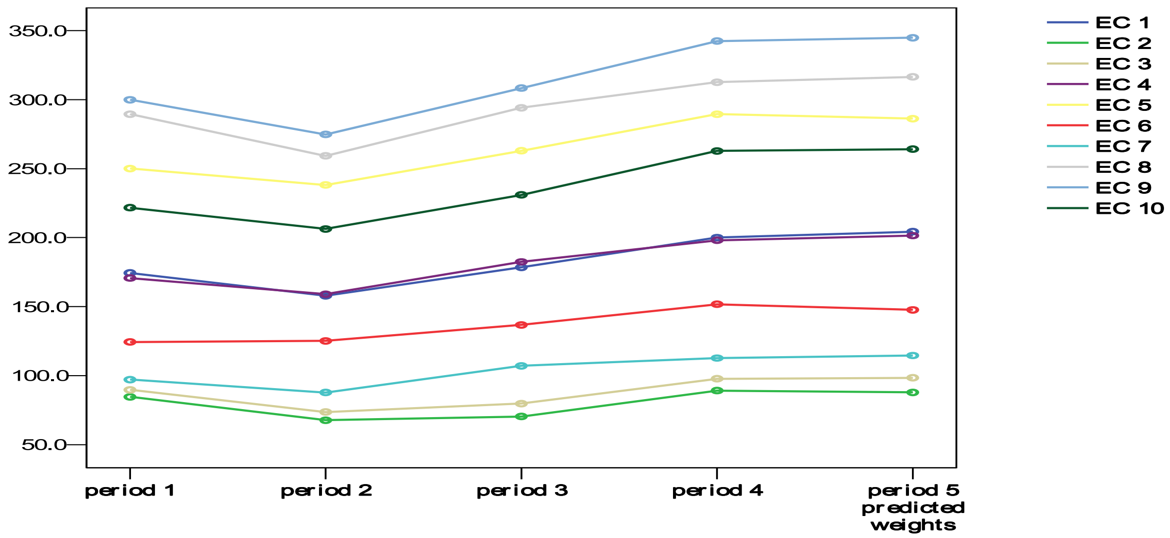

Fig. 2 The weight plot of periods for each CR.

In addition, the future trend of each EC to satisfy future CRs can be analysed in Table 6. According to the future trend of each EC, many ECs should be closely noticed since its importance has increased and could become the most important ECs to satisfy CRs in the future. On the other hand, the importance of EC2, EC5, EC6 have declined in the future. Different ECs should be considered differently for the computer designers and manufacturers with the much more analysed information.

The data mining cycle emphasizes the dataset information by repeating interaction activities. Since CRs can change rapidly, the database of CRs must be updated continually; therefore, the time series-based data mining cycle, proposed in this study, will continually update the database and continually identify the future CRs for the computer designers and manufacturers. These revised ECs will exactly satisfy with CRs, allowing the computer designers and manufacturers to the latest CRs, thus facilitating advanced design for green computers.

Table 6 Future trend of ten ECs for dynamic forecasting of life cycle

	EC1	EC2	EC3	EC4	EC5	EC6	EC7	EC8	EC9	EC10
Period 1	174.4	84.6	89.7	170.7	250.1	124.4	97.1	289.5	300	221.7
Period 2	157.9	67.7	73.6	159	238.2	125.2	87.8	259.3	274.8	206.3
Period 3	178.5	70.4	79.7	182.5	263	136.8	107.1	294.2	308.4	231
Period 4	200.1	89.1	97.7	198.1	289.5	151.7	112.7	312.7	342.5	262.9
Period 5 predicted weights	204.4	87.9	98.4	201.5	286.3	147.7	114.6	316.6	344.9	264.2

5 Conclusions

When considering green design, designers should incorporate the voices from the customers and because they are the driving force. This study uses data mining cycle in QFD to forecast future CRs for green design of life cycle. In addition, the time series-based data mining cycle can be applied to predict the weights and determine the trend of each CR for the next period. The use of time series-based data mining cycle to predict the weights is advantageous because it can (1) find the future trend of CRs; (2) provide the computer designers and manufacturers with reference points to satisfy CRs in advance. The results of this study can provide an effective procedure of identifying the trends of CRs and enhance dynamic forecasting of life cycle under green supply chain in the computer marketplace.

References:

[1] A. Ansari, B. Modaress, Quality function deployment: the role of suppliers. *International Journal of Purchase Management*, Vol. 30, No. 4, 1994, pp.28-35.

[2] A. Griffin, Evaluating QFD's use in US firms as a process for developing products, *Journal of Product Innovation Management*, Vol. 9, 1992, pp.171-187.

[3] C. F. Chien, A. Hsiao, I. Wang, Constructing semiconductor manufacturing performance indexes and applying data mining for manufacturing data analysis, *Journal of the Chinese Institute of Industrial Engineers*, Vol.21, 2004, pp.313-327.

[4] C. X. Feng, X. Wang, Development of empirical models for surface roughness prediction in finish turning, *International Journal of Advanced Manufacturing Technology*, Issue 20, 2002, pp.348-356.

[5] C. C. Shen, H. M. Chung, A study on the applications of data mining techniques to enhance customer lifetime value, *WSEAS TRANSACTIONS on INFORMATION SCIENCE and APPLICATIONS*, Issue 2, Vol.6, 2009, pp. 319-328.

[6] D. Pyle, *Data Preparation for Data Mining*, Morgan Kaufmann, California, 1999.

[7] D. Y. Sha, C. H. Liu, Using data mining for due date assignment in a dynamic job shop environment, *International Journal of Advanced Manufacturing Technology*, Issue 25, 2005, pp.1164-1174.

[8] E. Comparini and J. Cagan, *Reverse Engineering for Green Design of Products*, Carnegie Mellon, 1998.

[9] F. Consoli, D. Allen, I. Boustead, J. Fava, W. Franklin, A. A. Jensen, N. Deoude, R. Parrish, R. Perriman, D. Postlethwaite, B. Quay, J. Seguin, B. Vigon, *Guidelines for Life-Cycle Assessment: A Code of Practice*, Society of Environmental Toxicology and Chemistry, SETAC, 1993.

[10] G. S. Wasserman, How to prioritise design requirements during the QFD planning process. *IIE Transaction*, Issue 25, Vol.3, 1993, pp.59-65.

[11] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, *Time series analysis – Forecasting and control*. 3rd ed. Prentice Hall, Englewood Cliffs, NJ, USA, 1994.

[12] H. C. Harvey, *Time series models*, Halstead Press, New York, NY, USA, 1981.

[13] H. Min, A. Emam, Developing the profiles of truck drivers: a data mining approach, *International Journal of Physical Distribution & Logistics Management*, Issue 33, 2003, pp.149-162.

[14] I. Becerra-Fernandez, S. H. Zanakis, S. alczak, Knowledge discovery techniques for predicting country investment risk. *Computer and Industrial Engineering*, Issue 43, 2002, pp.787-800.

[15] J. R. Hauser, D. Clausing, *The house of quality*. Harvard Business Review, Issue 66, Vol.3, 1988, pp.63-73

[16] K. Keys, Design for Manufacture; Design for the life-cycle systems; life-cycle engineering, *IEEE International Electronic Manufacturing Technology Symposium*, 1988, pp. 62-72,

[17] K. W. Wong, S. Zhou, Q. Yang, M. S. Yeung, Mining customer value: from association rules to direct marketing, *Data Mining and Knowledge Discovery*, Issue 11, 2005, pp.57-79.

[18] L. P. Sullivan, *Quality function deployment (QFD): the beginning, the end, and the problem*. American Supplier Institute, 1987.

[19] M. Berry, G. Linoff, *Data Mining Techniques: for Marketing, Sales, and Customer Support*, Wiley, New York, 1997.

[20] Morro do Lena – Leiria, Campus de Azurém – Guimarães, Considering application domain ontologies for data mining, *WSEAS TRANSACTIONS on INFORMATION SCIENCE and APPLICATIONS*, Issue 6, Vol.9, 2009, pp. 1478-1492.

[21] M. Maddour, M. Elloumi, A data mining approach based on machine learning techniques to classify biological sequences, *Knowledge-Based Systems*, Issue 15, 2002, pp.217-223.

[22] P. Giudici, *Applied Data Mining: Statistical Methods for Business and Industry*, Wiley, England, 2003.

[23] R. M. Warner, *Spectral analysis of time - series data*. Guilford Press, New York, USA, 1998.

[24] R. Zust and R. Wagner, An approach to the identification and quantification of environmental effects during product life, *Annals of CIPR*, Vol. 41, No. 1, 1992, pp. 473-476.

[25] S. K. Dask, P. Yedlarajiah, R. Narendra, An approach for estimating the end-of-life product disassembly effort and cost," *International Journal of Production Research*, Vol. 38, No. 3, 2000, pp. 657-673,.

[26] S. G. Lee, S. W. Lye, M. K. Khoo, "A multi-objective methodology for evaluating product end-of-life options and disassembly,

International Journal of Advanced Manufacturing Technology, Vol. 18, 2001, pp. 148-1561.

[27] S. C. Wheelwright, K. B. Clark, *Revolutionizing product development: Quantum leaps in speed, efficiency, and quality*. New York: Free Press, 1992.

[28] T. C. Kuo, S. H. Chang, and S. H. Huang, "Fuzzy Multi-Attribute Decision-Making in Green Engineering," Proc. of the 5th International Conference on Industrial Engineering – Theory, Applications and Practice, Taipei, Taiwan, No. 278, 2000.

[29] T. C. Kuo, H. H. Wu, As customers green products development by applying grey relational analysis and green quality function deployment, *International Journal of Fuzzy Systems,* Vol. 5, No. 4, 2003.

[30] Y. M. Chas, S. H. Ho, K. W. Cho, D. H. Lee, S. H. Ji, Data mining approach to policy analysis in a health insurance domain, *International Journal of Medical Informatics*, Issue 62, 2001, pp.103-111.

[31] Y. Yu, K. Jin, H. C. Zhang, F. F. Ling, D. Barnes, A decision-making model for materials management of end-of-life electronic products," *Journal of Manufacturing Systems*, Vol. 19, No. 2, 2000, pp. 94-107.