# MODELS FOR ESTIMATING DESIGN EFFORT

**HAMDI A. BASHIR**

Department of Mechanical Engineering

McGill University

May 2000

A thesis submitted to the Faculty of Graduate Studies and Research in

partial fulfillment of the requirements for the degree of

Doctor of Philosophy

0-612-64508-8

# Abstract

In today's competitive environment, it is necessary to deliver products on time and within budget. Unfortunately, design projects have been plagued by severe cost and schedule overruns. This problem persists in spite of the significant advances that have been made in design technology over the last two decades. In most of the cases, the problem of overruns is due to poor estimation. The search for a solution has become even more pressing in the present era of shrinking product cycle times.

Driven primarily by this need, this thesis presents new effort estimation models. Unlike existing estimation techniques that are based on work breakdown structures with respect to process or product, the proposed models are based on a new metric for estimating product complexity, which is based on product functional decomposition. The validity of the metric as a good predictor of design effort was tested using data obtained from an experiment involving simple design tasks, and empirically using historical data collected for 32 projects from 3 companies.

The performance of the new effort estimation models was tested in terms of a number of objective criteria. The results indicated that the average estimation error of the models ranged from 12% to 15%. The improvement in estimation accuracy accomplished by the models ranged from 52% to 64% compared to estimates originally made by the companies which had errors from 27% to 41%.

Moreover, models for estimating cost and duration, as well as updating the estimates during project execution, were derived. The applications of the derived models are described through demonstrative examples. Thus, a complete methodology is given for the estimation of project effort and duration.

# Résumé


Dans l'environnement concurrentiel d'aujourd'hui, il est nécessaire de livrer les produits aux clients dans le temps et selon le budget établi. C'est malheureusement, la difficulté principale à laquelle se heurtent les projects de conception. Ce problème persiste malgré les avancées significatives de la technologie de conception au cours des deux dernières décennies. Dans la plupart des cas, la cause du problème est la mauvaise estmation des besoins. La recherche d'une solution est devenue de plus en plus pressante à cause de la tendance généralisée de la réduction du cycle de développement de produits.

Guidée avant tout par ce besoin, cette thèse propose de nouveaux modèles d'estimation. À la difference des techniques d'estimation existantes qui sont basées sur le fractionnement du processus ou du produit, les modèles proposés sont orientés sur une nouvelle mesure de performance définie dans l'optique d'estimation de la complexité d'un produit. Ce modèle est orienté sur la décomposition du produit par analyse fonctionnelle. Cette mesure a été validée en tant que prédiction d'effort requis pour la conception, à travers de expériences réalisées dans le cadre d'activités simples de conception, et avec une analyse empirique utilisant des données provenant de 32 projets réalisés dans trois compagnies différentes.

La performance de ces nouveaux modèles d'estimation d'effort a été examinée en termes de critères objectifs multiples. Les résultats ont indiqué que la moyenne de l'erreur d'estimation des modèles était entre 12% et 15%. L'amélioration dans la précision accomplie par les modèles se situe entre 52% et 64%, par rapport aux estimations originales des compagnies qui avaient des erreurs entre 27% et 41%.

De plus, des modéles d'estimation des coûts et des délais, ainsi qu'une mise à jour pendent l'exécution du projet, ont été développés. L'application de ces modèles est illustrée dans la thèse à l'aide d'examples pratiques. Cet ouvrage présente donc une méthodologie complète d'estimation d'effort et de durée d'un projet de conception.

# Acknowledgements

Humble thanks are first offered to God

I wish to express my deepest gratitude and sincere appreciation to my supervisor Professor Vince Thomson, for his academic as well as financial support, and his invaluable and constructive guidance throughout all the stages of this study.

I am deeply indebted to the University of Garyounis and the Ministry of Higher Education of Libya for their sponsorship.

My thanks to all those at the Canadian Bureau for International Education (CBIE) for administering the funds provided by the Ministry of Higher Education of Libya.

I wish to express my special thanks to Paul Brooks of Bombardier Inc., who spent much of his time in evaluating the design tasks for the experiment.

My thanks are also extended to the design managers at Nortel Networks, Canadian Marconi Company, and General Electric Hydro who provided the data, which were used for constructing and validating the models described in this thesis.

Last but not least, I would like to dedicate this thesis to my father and mother who have been a continuous source of love and encouragement throughout my whole life.

# Table of Contents

## CHAPTER 6: ARTIFICIAL NEURAL NETWORK MODELS



## CHAPTER 7: AN ANALOGY-BASED MODEL



## CHAPTER 8: APPLICATIONS OF THE MODELS



## CHAPTER 9: CONCLUSIONS





## APPENDIX I: PROJECT DATA

## Appendix II: Design Tasks



## Appendix III: Use of the Principal Right Eigenvector in the Analogy-Based Model



## Appendix IV: Analogy-Based Model Computations

# Nomenclature

| | |
|---|---|
| $A$ | = *pairwise comparison matrix* |
| $A_i$ | = *activity i* |
| $AC$ | = *the average cost per hour* |
| $a$ | = *constant* |
| $b$ | = *constant* |
| $C_k$ | = *employee category k in the context of activity i* |
| $CI$ | = *consistency index* |
| $CM$ | = *consistency measure* |
| $CT$ | = *concept to customer* |
| $CF_{i,k}$ | = *the relative complexity of an activity in the context of a particular building block compared to an average complexity of the same type in the same environment* |
| $c_n$ | = *constant* |
| $D_m$ | = *effort driver (factor m)* |
| $DE$ | = *difficulty to expertise ratio* |
| $DT$ | = *development time* |
| $DP$ | = *involvement of design partners* |
| $dr_i$ | = *the difference in ranks of the* $i^{th}$ *pair of data* |
| $E$ | = *actual design effort* |
| $E_i$ | = *actual design effort of project i* |
| $E_r$ | = *the input of reference project r* |
| $E_u$ | = *the estimated design effort of upcoming project u* |
| $E_{ur}$ | = *the estimated design effort of upcoming project u using reference project r* |
| $ES_i$ | = *effective size* |
| $\hat{E}$ | = *estimated design effort* |
| $\hat{E}_b$ | = *the required design effort for a building block* |

| | |
|---|---|
| $\hat{E}_e$ | = *estimated engineering effort* |
| $\hat{E}_i$ | = *estimated design effort of project i* |
| $\overline{E}_i$ | = *the mean of the values* $E_i$ |
| $e$ | = *the base of the natural logarithm* |
| $F_j$ | = *number of functions at level j* |
| $FP$ | = *use of a formal process* |
| $FT$ | = *use of cross-functional teams* |
| $g$ | = *number of groupings of different tied ranks* |
| $H_o$ | = *null hypothesis* |
| $H_1$ | = *alternative hypothesis* |
| $KW$ | = *Kruskal-Wallis test* |
| $k$ | = *number of treatments* |
| $l$ | = *number of levels* |
| $M_{rf}$ | = *a multiplier which adjusts the productivity of reference project r due to the influence of factor f* |
| $MMRE$ | = *the mean magnitude of relative error* |
| $m$ | = *number of influencing factors* |
| $N$ | = *number of projects* |
| $N_T$ | = *total number of observations* |
| $NE$ | = *number of elements of an individual fundamental circuit in the context of activity i* |
| $NN$ | = *newness (percentage of change)* |
| $NRECP(A_i, C_k)$ | = *the partial activity effort of employee category k in the context of activity i* |
| $n$ | = *sample size* |
| $n_j$ | = *number of observations in the* $j^{th}$ *treatment* |
| $nr$ | = *number of reference projects* |
| $ns$ | = *number of subsamples* |
| $O_r$ | = *the output of reference project r* |
| $O_u$ | = *the output of upcoming project u* |
| $P_r$ | = *the productivity of reference project r* |

| | |
|---|---|
| $P_{ur}$ | = *the estimated productivity of upcoming project u using reference project r* |
| $Ps_i$ | = *pseudo-value for the entire sample omitting subsample i* |
| $PC$ | = *product complexity* |
| $PC_i$ | = *initial product complexity* |
| $PC_n$ | = *new product complexity* |
| $PF_{i,k}$ | = *productivity factor of employees in category k involved in activity i* |
| $PRED(l)$ | = *prediction at a given level l* |
| $p$ | = *level of significance* |
| $R$ | = *the computed CI of randomly generated matrices* |
| $R_j$ | = *average of the ranks in the* $j^{th}$ *treatment* |
| $R^2$ | = *the coefficient of multiple determination* |
| $RE$ | = *number of repeated elements* |
| $RW$ | = *time to be spent on rework* |
| $RF_{i,k}$ | = *reuse factor which captures the reduction in effort due to the reuse of entire building blocks* |
| $r_s$ | = *Spearman rank-order correlation coefficient* |
| $S$ | = *team size* |
| $SR$ | = *severity of requirements* |
| $s$ | = *matrix size* |
| $T$ | = *total direct manpower cost* |
| $TD$ | = *type of drawings submitted to the customer* |
| $TT$ | = *total time* |
| $t$ | = *time* |
| $t_d$ | = *project duration* |
| $t_{dn}$ | = *new estimate for project duration* |
| $t_i$ | = *number of tied ranks in the* $i^{th}$ *grouping* |
| $t_o$ | = *the time at which peak effort occurs* |
| $v$ | = *number of variables* |
| $WF_i$ | = *a weighting factor which accounts for the impact of the number of repeated elements in the context of activity i* |

| | |
|---|---|
| $w$ | = *the principal right eigenvector* |
| $w_k$ | = *weight assigned to function k* |
| $w_{rf}$ | = *the extracted weight corresponding to reference project r and influence factor f* |
| $w_{uf}$ | = *the extracted weight corresponding to upcoming project u and influence factor f* |
| $X$ | = *independent vector* |
| $x$ | = *the average ratio of actual design effort spent during the last month to the actual total design effort* |
| $\hat{Y}(X)$ | = *expected value* |
| $y$ | = *cumulative manpower* |
| $y_o$ | = *random variable* |
| $y'$ | = *manpower required in time period t* |
| $\alpha$ | = *shape parameter* |
| $\alpha_n$ | = *new shape parameter* |
| $\hat{\beta}$ | = *the least squares estimator of the entire sample* |
| $\tilde{\beta}$ | = *the jackknife estimator* |
| $\hat{\beta}_{-i}$ | = *the least squares estimator of the entire sample omitting subsample i* |
| $\gamma$ | = *the average ratio of actual design effort to time-to-peak* |
| $\lambda_{\max}$ | = *the largest eigenvalue* |
| $\sigma$ | = *smoothing factor* |
| $\chi^2_{.01,2}$ | = *the chi-square distribution with a degree of freedom of 2 at a level of significance of .01* |

# List of Figures

# List of Tables

# Chapter 1

# Introduction

*"When you can measure what you are speaking about, and express it in numbers, you know some thing about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind"*

*Lord Kelvin, Popular Lectures and Addresses, 1889 (Cook, 1982)*

## 1.1 Motivation

In many aspects, design projects are not very different from projects in any other discipline; they all require management skills, i.e., the ability to plan, organize, coordinate, and control. However, design projects are characterized by a lack of easily identifiable and measurable items that can provide data for the estimation of effort and feedback on performance. Schedule slippage and cost overrun are typical for most design projects. According to Bounds (1998), only 26% of the projects in the United States are completed on time and within budget. In published papers, the reported average schedule overrun ranges from 41% to 258%, and cost overrun ranges from 97% to 151% (Norris, 1971; Murmann 1994).

### 1.1.1 Consequences of Overruns

Overruns have many consequences, such as:

- In some situations, cost or schedule overruns lead to project termination (Bronikowski, 1986).
- Schedule overrun increases the risk of product obsolescence due to the increased risk of missing the market window (Leech, 1972; Bronikowski, 1986; Cordero, 1991). According to (Evans, 1990), delay is deadly, and in many cases can lead to project

failure. In the auto industry, one study indicates that each day of delay costs an automobile firm over US$1 million in lost profits (Clark et al., 1987).

- An initial delay in a project can engender further delays. Once a project falls behind schedule, one or both of the following measures are usually taken: extending the working hours of staff or increasing the number of people on the project. The first measure may increase the stress on the team and lead to an increase in the error rate. For instance, DeMarco (1982) points out, "people under time pressure don't work better, they just work faster." As a result, the amount of rework may be increased and the completion time may be extended. The second measure requires additional communication, caused by added personnel, which usually exacerbates the situation. Brooks (1975) states, "The natural response to a late project is to add manpower, like dousing a fire with gasoline. This makes matters worse, much worse." The ability to add manpower is limited. Moving experienced people from one project to another just endangers the 'robbed' project. Also it takes up to six months to get new hires up to speed on a project; this exacerbates the communication task of team members.

### 1.1.2 Causes for Overruns

There are many candidate causes for overruns. In order to identify the major causes, two studies were conducted, one by Thamhain and Wilemon (1986) the other by Phan et al. (1988).

#### 1.1.2.1 Study by Thamhain and Wilemon

In this study, data was collected mostly by questionnaires from 304 project leaders (general managers and project managers) of 183 technical projects. Those questioned had an average of 5.2 years experience in project management. The average project duration was 12 months, and an average of 8 people worked on a project. The survey investigated what the managers believed to be the reason for cost and schedule overruns. Their reasons were ranked in order of importance and the results are shown in Table 1.1. This table indicates that, for general managers, four of the top five reasons for overruns are specifically related to planning, while for project managers, the top five reasons are related to planning and project dynamics. Furthermore, in spite of their disagreement on the relative importance of 9 of the 15 causes, general and project managers did strongly agree on one cause, unrealistic project planning.

**Table 1.1** Reasons for schedule and cost overruns (Thamhain and Wilemon, 1986)

| Rank by General Manager | Rank by Project Manager | Cause | Agreement between GM & PM |
|---|---|---|---|
| 1 | 10 | Insufficient front-end planning | Disagree |
| 2 | 3 | Unrealistic project plan | Strongly agree |
| 3 | 8 | Project scope underestimated | Disagree |
| 4 | 1 | Customer/ management changes | Disagree |
| 5 | 14 | Insufficient contingency planning | Disagree |
| 6 | 13 | Inability to track progress | Disagree |
| 7 | 5 | Inability to detect problems early | Agree |
| 8 | 9 | Insufficient number of check points | Agree |
| 9 | 4 | Staffing problems | Disagree |
| 10 | 2 | Technical complexity | Disagree |
| 11 | 6 | Priority shifts | Disagree |
| 12 | 10 | No commitment by personnel to plan | Agree |
| 13 | 12 | Uncooperative support group | Agree |
| 14 | 7 | Sinking team spirit | Disagree |
| 15 | 15 | Unqualified project personnel | Agree |

#### 1.1.2.2 Study by Phan et al.

In this study, questionnaires were sent to 827 members of the American Institute of Certification of Computer Professionals. The 191 respondents were involved in projects with an average duration of 14 months and an average of 17 people working on a project. The cause of overruns was one of the points addressed. Forty four percent of the respondents indicated that over optimistic planning was the usual cause of overruns. Minor changes, major changes and the lack of tools were given as a cause by 33%, 36% and 17% of the respondents respectively.

While this survey was confined to information system development projects, the author believes that the characteristics of the design process are similar regardless of the object of design, i.e., mechanical, electrical, software; thus, lessons learned in one design discipline should shed some light on problems in another. This idea is demonstrated by the work in this thesis.

It can be concluded from the above two surveys that poor estimation of cost and duration is one of the major causes of project overruns. The problem of poor estimation is most likely due to inherent weaknesses in available approaches, which make them ineffective in producing realistic estimates. The existence of such a problem in the present era of shrinking product cycle times has made the need for sound estimators more acute than ever before. Improving estimation accuracy is a vital issue not only for companies that use traditional design approaches, but also for those adopting newer approaches such as concurrent engineering. In other words, reducing the cycle time of a project is a futile effort without being better able to estimate the required time within an acceptable degree of error, and thereby, reduce the probability of overruns in time and cost. The emphasis must be on improving the accuracy of estimating design effort. This is because of the following.

- Since labor costs make up the majority of the cost for most design projects, effort estimation can provide a good estimate of project cost.
- Scheduling cannot be made without determining the available resources and estimating the required resources (effort).
- Without good estimates of project duration and cost, there is no way of subsequently determining if a project is on schedule or within budget.

In other words, reliable estimation of design effort is a necessary prerequisite for developing reliable schedule and cost estimates, as well as for monitoring the progress of a project (Adrangi and Harrison, 1987).

## 1.2 Scope and Objectives

This research is limited to design projects. A design project is a combination of interrelated activities that must be executed in a particular order to complete a task (Elsayed and Boucher, 1985). The task involves converting an idea or market need into detailed information from which a product or a system can be produced (Hales, 1987). The projects studied during the research for this thesis were restricted to mechanical and electronic design; however, the intent was to develop a methodology which would be applicable to any design project.

The research objectives are:

1) To develop a simple and useful metric for estimating product complexity which can be used to estimate the required design effort.
2) To develop models for estimating design effort. The models attempt to be:
   - applicable to a wide range of design projects
   - reasonably accurate[1]
   - easy to use
   - parsimonious
3) To compare design effort estimation models that use traditional tools, e.g., regression analysis, and new tools, e.g., artificial neural networks.

While its focus is on developing models for estimating design effort, this research has an additional objective, which is to derive models for estimating project cost and duration, as well as being able to updating these estimates during project execution.

## 1.3 Thesis Organization

Review of the available relevant literature is presented in Chapter 2. The methodology applied for developing the models proposed in this thesis including data collection and the criteria used for evaluating the performance of the models are the subject of Chapter 3. Chapter 4 proposes a new metric for estimating product complexity and describes the validation methods used. Parametric estimation models using traditional regression analysis are presented in Chapter 5. Chapter 6 introduces artificial neural networks as a promising tool for estimating design effort. Complementary to the models described in Chapters 5

---

[1] The average estimation error to be not more than 25%, and 75% of the estimations to be within 25% of the actual values.

and 6, an analogy-based model is proposed in Chapter 7. Potential applications of the models are demonstrated in Chapter 8. Lastly, Chapter 9 concludes the thesis with a summary of the findings, identifies the limitations, and makes suggestions for future research.

## 1.4 Contributions of this Thesis

This thesis claims the following contributions:

- The development of a very effective metric for estimating product complexity. The metric is based on product functional decomposition. Although a significant body of work does exist on functional decomposition, no previous research has dealt with quantifying product complexity in terms of functional decomposition.
- The development of new models for estimating design effort. These models not only improve the accuracy of effort estimation, but also make answering the following type of questions easier:
  - How much will a project cost?
  - How long will a project take?
  - What will happen if changes in requirements are made?
- Because they are based on product functionality, the developed models have the potential for being widely applicable in many disciplines.
- This is the first study that addresses a novel application of neural networks. As of writing of this thesis, no researchers had yet applied this technology to this specific discipline.
- Adaptation of Norden's model (effort distribution versus time) to estimate the duration of design projects and to model the variation of project duration with changing product requirements and/or staff levels.

It is worth mentioning that most of the materials presented in this thesis have been published or accepted for publication as follows:

Bashir, H. A., V. Thomson, 1999a, Metrics for Design Projects: A Review, *Design Studies*, Vol. 20, No. 3, pp. 263-277.

Bashir, H. A., V. Thomson, 1999b, Estimating Design Complexity, *Journal of Engineering Design*, Vol. 10, No. 3, pp. 247-257.

Bashir, H. A., V. Thomson, 1999c, Models for Estimating Design Effort and Time, *Design Studies*, (in press).

Bashir, H. A., V. Thomson, 1999d, An Analogy-based Model for Estimating Design Effort, *Design Studies*, (in press).

Bashir, H. A., V. Thomson, 1999e, Estimating Design Effort Using Artificial Neural Networks, *Proceedings of the 3rd International Conference on Engineering Design and Automation*, Vancouver, Canada, pp. 344-351.

Bashir, H. A., V. Thomson, 1999f, A Quantitative Estimation Methodology for Design Projects, *Proceedings of the International Conference on Industrial Engineering and Production Management (IEPM'99)*, Glasgow, U.K. pp. 498-506.

# Chapter 2

# Review of Literature

The main conclusion of the review of the work in Chapter 1 is that the inadequacy of the available estimation methods is one of the major factors contributing to the problem of design project overruns, and that there is a need for constructing new models for estimating design effort. Before describing these models, this chapter reviews the existing estimation methods, which mainly fall under one of the two following categories: expert judgement and the metrics approach. This is a brief review of the work done to date in this area.

## 2.1 Expert Judgement

In its simplest form, expert judgment involves consulting one or more estimators who use their experience from past projects to arrive at an estimate. Since the late 1940's, a number of structured expert judgement methods have been proposed. These methods include:

- Delphi Technique
- Critical Path Method (CPM)
- Program Evaluation and Review Technique (PERT)
- Work Breakdown Structure (WBS)

### 2.1.1 Delphi Technique

This technique was developed by the Rand Corporation in 1948 (Helmer, 1966). A group of experts is asked to make individual predictions secretly. The average estimate is calculated and presented to the group. The experts are then given the opportunity to revise their estimates, if they so wish. The process is repeated until none of the experts want to change his or her estimates any further.

### 2.1.2 Critical Path Method (CPM)

CPM was developed by Kelley and Walker (1959). It uses expert judgement to provide duration estimates for project activities, which are arranged in a directed graph. Then, the total estimated time of all the activities on the longest path is considered as the total duration of the project and the summation of the estimated costs of all the activities is considered as the project cost.

### 2.1.3 Program Evaluation and Review Technique (PERT)

PERT was developed in 1959 as a joint effort of Booz, Allen and Hamilton and the U.S. Navy's Special Project Office. The technique is very similar to CPM, except it allows for uncertainty in the time estimates of activities (Levin and Kirkpatrick, 1966).

### 2.1.4 Work Breakdown Structure (WBS)

WBS has been introduced by the United States Department of Defense (DOD) in 1963 and applied widely for schedules and cost estimation. The technique can be either product or process oriented. In the former, the end product is broken down into subsystems. These subsystems are further subdivided into sub-subsystems, and so on. While in the latter, the end product is broken down into the processes required to produce it (Mansuy, 1991).

The use of pure expert judgement in any form (simplest or structured) has often led to unsatisfactory results. This is because of the following:

- The accuracy of estimation depends on the competence, experience, objectivity, and perception of the estimator. Experts estimate using analogies with other projects. However, projects that appear to be similar can in fact be quite different. Even when it is known how one project differs from another, it is not always apparent how the differences affect cost and time. This also means that sensitivity analysis on such estimations is not easily performed (Conte et al., 1986).
- Because they are usually done by people who are involved in the project, the estimation can be biased (Conte et al., 1986; Adrangi and Harrison, 1987). According to DeMarco (1982), people underestimate the time they themselves will take to do something.

- The benefit of the structured methods is very limited, especially for large projects (Leech, 1972; Putnam, 1987b). For example, Kingel (1966) ascertained that under many circumstances, PERT calculations could be biased, and thus, give poor estimates.

## 2.2 The Metrics Approach

In this approach, the use of subjective estimation is minimized by assigning quantitative indices to the attributes of project entities, e.g., design complexity, technical difficulty, design team experience, etc. These indices are used to construct estimation models. A metrics approach has the potential of allowing managers to estimate design effort and duration more accurately, and to monitor the development of a product more objectively (Bashir and Thomson, 1999a). This is based on the following.

- A metrics approach is a more systematic way of overcoming the problem of biased estimation that characterizes most of the available estimation techniques.
- Measurement is extremely important in managing any process. "If a process is not being measured, then it is not being managed" (Rummer and Brache, 1990).
- It has emerged as an effective management tool in disciplines such as software development. According to DeMarco (1982), companies that use software metrics produce substantially better estimates of effort and duration. Also, the application of software metrics has proven to be effective in improving software quality and productivity (Moller and Paulish, 1993).

On the other hand, the metrics approach does not take into account unusual situations or a changing environment; therefore, it is only useful in a relatively constant environment. The following sections review studies that have adopted the metrics approach to address the following type of questions.

- How much effort will be required?
- How many people are needed at any one time?
- How long will it take?

The studies reviewed here are by Norden (1964, 1970), Griffin (1993, 1997), and Jacome and Lapinskii (1997).

### 2.2.1 Study by Norden

Norden (1964, 1970) noted that there are regular patterns of manpower increase and decrease independent of the type of work done, that is related to the way people solve problems. On the basis of his analysis, Norden succeeded in creating a useful model (equation (2.1)) that describes the utilization of manpower during each of the design phases: planning, design, model, and release. Depending on the amount of overlap between the phases, the entire project cycle may be represented or at least approximated by equation (2.1).

$$y' = 2\hat{E}\alpha t e^{-\alpha t^2} \tag{2.1}$$

where:

$y$ = *manpower in appropriate units, e.g., hours or man-months*

$y'$ = *manpower in appropriate units, e.g., hours or man-months, required in time period t*

$\hat{E}$ = *total estimated design effort stated in the same units as y, e.g., hours or man-months*

$\alpha$ = *a shape parameter defined by the point in time at which* $y'$ *reaches its maximum value*

$t$ = *time, in equal units such as weeks or months*

$e$ = *the base of the natural logarithm*

The shape parameter, $\alpha$, is computed as follows (Norden, 1964, 1970).

At time, $t_o$, at which peak effort occurs:

$$\frac{dy'}{dt_o} = 2\hat{E}\alpha e^{-\alpha t_o^2} - 4\hat{E}\alpha^2 t_o^2\, e^{-\alpha t_o^2} = 0$$

$$2\hat{E}\alpha e^{-\alpha t_o^2} = 4\hat{E}\alpha^2 t_o^2\, e^{-\alpha t_o^2}$$

$$\alpha = \frac{0.5}{t_o^2} \tag{2.2}$$

Note that the integral of equation (2.1) is:

$$y = \hat{E}(1 - e^{-\alpha t^2}) \tag{2.3}$$

where:

$y$ = *the cumulative manpower used through time $t$, stated in the same units as $\hat{E}$, e.g., hours or man-months*

Norden's model defines the relationship between effort and duration, i.e., if the effort of doing a design can be determined, then, the duration of a project can be predicted. The shape of the curve defined by the model is shown in Figure 2.1 for manpower usage versus time, where the area under each curve is total effort for a design phase and the time from the start to a point at the end of the curve[2] is phase duration.

One of the shortcomings of the model is the need to subjectively estimate its two fundamental parameters: total design effort, $\hat{E}$, and the shape parameter, $\alpha$.

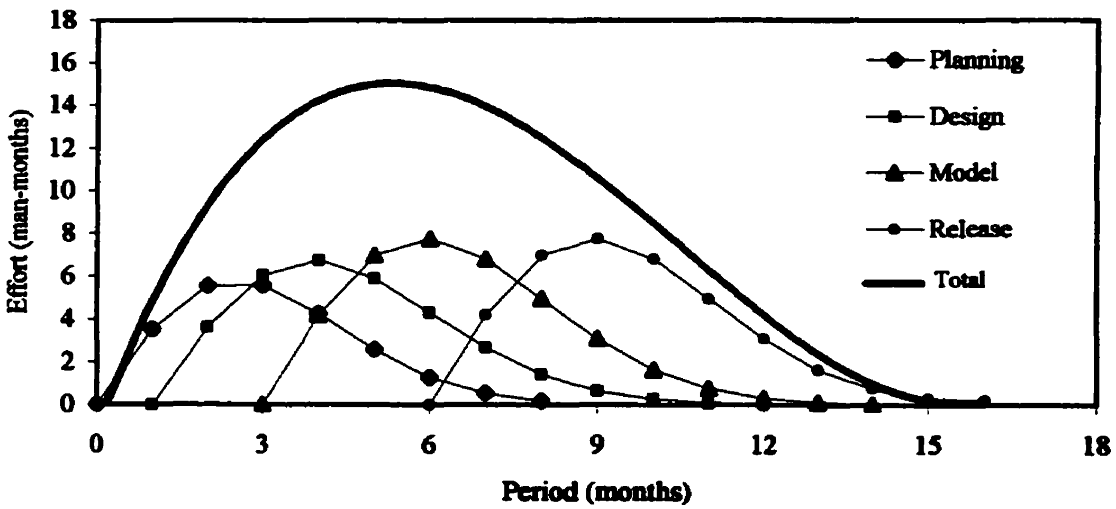



**Figure 2.1** Typical manpower pattern for a hypothetical project

[2] Since the curve tails out to infinity, a method for estimating an end point is presented in Chapter 8.

### 2.2.2 Study by Griffin

Griffin (1993) has introduced a number of metrics and classified them under three categories: project characteristics, outcome, and development process metrics. Metrics for project characteristics include complexity and amount of change. Outcome metrics include time through each phase (introduction, development time, concept to customer, total time), cost of development, product commercial success, and customer satisfaction. Metrics for the development process include type of process used, delivery of customer needs, and others. Most importantly, Griffin investigated the possibility of establishing useful relationships between development time (the time between the first development team meeting and the date of first product for sale) and the following:

**Project complexity:** The number of functions that the product performs (product complexity) and the number of technologies or functional specialties involved (management complexity).

**Amount of change:** The percentage of change that has been introduced in the product and the manufacturing process with respect to the previous generation.

**Use of a formal process:** A formal process is usually called a phase review or stage-gate process. In this process, the development is divided into a series of phases, and at the end of each phase, completed activities are reviewed and approved.

To do this study, historical data were collected for 45 projects. The following summarizes the results of the analysis:

- The relationship between development time and the amount of change was positive and statistically significant[3] as indicted by the following regression model[4]:

$$DT = 5.3 + 0.18\, NN \qquad (2.4)$$

---

[3] The coefficient was at $p < 0.01$ level of significance.

[4] This model is based on projects of similar complexity across one company (nine data points), where no formal process was used. For more details, refer to Griffin (1993).

where:

$DT$ = *development time in months (the time between the first development team meeting and the date of first product for sale)*

$NN$ = *percentage change from previous product (ranges from 0% to 100%)*

- Complexity had an influence on product development time. However, due to limitations in the sample data, no relationship was deduced between them.
- Cycle time was more predictable if a formal process for engineering was used by the development team.

As an extension to the above work, Griffin (1997) published a recent study in which she developed the following multivariable models:

$$DT = 8.4 + 6.1\,PC + 0.18\,NN - 1.9\,FP - 0.09\,FT \tag{2.5}$$

$$CT = 10.4 + 3.8\,PC + 0.32\,NN + 0.1\,FP - 0.16\,FT \tag{2.6}$$

$$TT = 13.8 + 4.5\,PC + 0.30\,NN + 0.5\,FP - 0.15\,FT \tag{2.7}$$

where:

$DT$ = *development time in months (the time between the first development team meeting and the date of first product for sale)*

$CT$ = *concept to customer in months (the time between approval of strategy or idea and the date of first production)*

$TT$ = *total time in months (begins when the idea for the product first surfaces and ends with the date of first production)*

$PC$ = *product complexity (number of functions)*

$NN$ = *newness (percentage of change)*

$FP$ = *use of a formal process (dichotomous (0 – 1, no –yes))*

$FT$ = *use of cross-functional teams (dichotomous (0 – 1, no –yes))*

The above models were developed using data sets gathered for 343 projects from different companies; however, these models only account for a small portion of data variation ($R^2$ ranges from 0.15 to 0.30). Griffin argues that the heterogeneity of the projects is the only explanation for the unexplained variation. In addition, the author suggests the following other possible reasons:

- poor selection of independent variables,
- the assumption of model linearity, and/or
- the weakness of the product complexity metric, which will be discussed further in Chapter 4.

Moreover, Griffin assumes that as complexity increases or percentage change increases, the development time increases. This assumption is not always true; in fact, as the complexity increases or percentage change increases, the effort, but not necessarily the development time, increases. More complexity leads to more effort; development time depends on effort, resource availability, and on the amount of the work that can be done concurrently.

### 2.2.3 Study by Jacome and Lapinskii

To estimate the effort required for designing a new electronic product, Jacome and Lapinskii (1997) propose a process-oriented model which takes into account three major factors: size, complexity, and productivity. The first factor captures the size (number of gates or transistors) of the design objects to be considered in the design task. The second factor accounts for the task's relative difficulty in a particular environment. The third factor considers the rate (effort per gate or transistor) at which the task progresses.

In order to apply the model, the product is decomposed into manageable units (components) called building blocks. The required design effort for each building block is estimated by using equation (2.8).

$$\hat{E}_b = \sum_i \sum_k NRECP\ (A_i, C_k) \qquad (2.8)$$

where:

$\hat{E}_b$ = *the required design effort for a building block in man-months*

$A_i$ = *activity i (for example, architectural design)*

$C_k$ = *employee category k in the context of activity i (for example, architectural designers)*

$NRECP(A_i, C_k)$ = *the partial activity effort of employee category k in the context of activity i (the number of architectural-designer-months required for architectural design)*

$NRECP(A_i, C_k)$ is computed as follows:

$$NRECP(A_i, C_k) = ES_i \, CF_{i,k} \, PF_{i,k} \, (1 - RF_{i,k}) \quad (2.9)$$

where:

$ES_i$ = *effective size which is defined as the subset of the fundamental circuit types or the building block abstraction in which the activity i applies*

$CF_{i,k}$ = *the relative complexity of an activity in the context of a particular building block compared to an average complexity of the same type in the same environment*

$PF_{i,k}$ = *productivity factor (effort in man-months per gate or transistor) of employees in category k involved in activity i*

$RF_{i,k}$ = *reuse factor which captures the reduction in effort due to the reuse of entire building blocks*

The effective size of an individual fundamental circuit type in the context of activity *i* ($ES_i$) is given by equation (2.10).

$$ES_i = NE - (WF_i \, RE) \quad (2.10)$$

where:

$NE$ = *number of elements of an individual fundamental circuit in the context of activity i*

$WF_i$ = *a weighting factor which accounts for the impact of the number of repeated elements in the context of activity i*

$RE$ = *number of repeated elements*

One advantage of Jacome and Lapinskii's model is the use of a combination of two estimation approaches (a bottom-up and a metrics approach). Such a combination helps to avoid the weaknesses of any single approach and to capitalize on their joint strengths (Boehm, 1981). However, the formulation of the productivity factor, $PF_{i,k}$, by Jacome and Lapinskii works in the opposite way to the usual definition of productivity. $PF_{i,k}$ should be defined as number of gates or transistors per unit of effort, i.e., output/input. Consequently, equation (2.9) would be reformulated as follows:

$$NRECP(A_i, C_k) = ES_i\ CF_{i,k}\ \frac{(1 - RF_{i,k})}{PF_{i,k}} \tag{2.11}$$

The model developed by Jacome and Lapinskii is presently being used in a software system for effort estimation for electronic design.

## 2.3 Summary

In this chapter, the available estimation techniques have been outlined, and more details were given about the major studies that have adopted a metrics approach to provide essential project estimates. These estimates included design effort (Jacome and Lapinskii), design effort distribution with time (Norden), and duration (Griffin). It can be concluded from this review that Jacome and Lapinskii's model is relevant only for a specific application. Griffin's models do not appear to predict project duration well, and as pointed out, the fits of the models to actual project data are very poor.

Given that Norden's model can estimate the required manpower across the entire life of a project as well as project duration once effort is estimated; then, further research is clearly needed to be able to develop good effort estimation models for the general design process. Nevertheless, solace should be taken from the fact that good estimation models have been developed in certain domains, e.g., software development (Walston and Felix, 1977; Albecht, 1979; Boehm, 1981). It remains to be seen how effective new models can be for the general design process.

# Chapter 3

# Methodology

Estimation models can be classified into two major categories: empirical and theoretical. However, at these early investigative stages, only empirical models are possible. Empirical models are generally derived from historical data using the methodology summarized in Figure 3.1. As Figure 3.1 shows, this approach involves the gathering of data about project characteristics to identify the most significant factors to be included in the model. Once these factors are identified, a theory of their interaction is formulated, and a prototype for a model is proposed. Then, the model is evaluated using one or a combination of criteria, so that a determination can be made about whether the model is acceptable. If not, the theory is revised and a new model is proposed.

## 3.1 Data Collection

To build and test a model, data on past projects are required. The number of projects depends on the number of variables to be included in the model. Theoretically, three projects are sufficient for one variable model. Generally, $v + 2$ projects are needed for a model involving $v$ variables (Conte et al., 1986; Fenton and Pfleeger 1997). However, in order to detect the underlying relationships, the number of data points must be determined according to the variability of data; in other words, the more heterogeneous the projects are, the greater the number of projects that are required.

### 3.1.1 Source of the Data

The models described in this thesis were built and tested using historical data from 32 previously completed design projects from three different companies. These companies are:

- Nortel Networks (NN)

- Canadian Marconi Company (CMC)
- General Electric Hydro (GE)

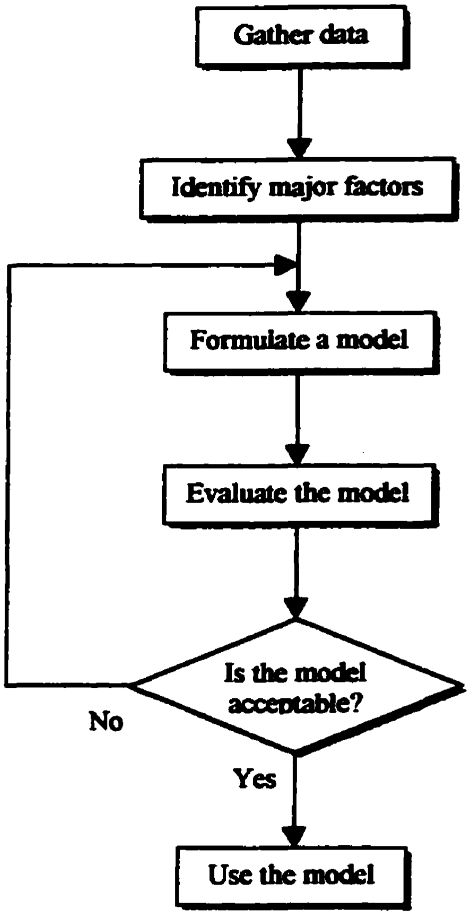



**Figure 3.1** Construction of a model

NN is recognized as a world leader in the design and manufacture of electronic and electrical products including wireless communication networks power systems, etc. Historical data from 5 previously completed projects related to the development of battery chargers were obtained from NN. These projects were carried out in the period between 1994-1998.

CMC is recognized as one of the world leaders in the design and manufacture of high-technology electronic products including avionics, communication systems, and others. Historical data from 12 previously completed projects related to the development of communication systems were obtained from CMC. These projects were carried out in the period between 1996-1998.

GE Hydro is a world leader in the design and construction of generators and turbines. Historical data from 15 previously completed projects related to the development of generators were obtained from GE. These projects were carried out in the period between 1985-1999.

To ensure consistency of data, one person supplied data for each company. Information provided to the companies included a glossary of terms, instructions for filling out forms, and some demonstrative examples. Interviews were also conducted to check the data. The collected data are presented in Appendix I.

Below are the assumptions and definitions underlying the use of the collected data.

1) Design managers honestly followed the given guidelines and instructions to provide the required data; in other words, the data were reasonably accurate, correct, and consistent.
2) Design effort[5] was the total time in hours or man-months spent by all the people involved directly in the project including design managers. In this thesis, a man-month consists of 152 hours of working time[6]. Since an estimate cannot be made until design requirements are determined, the models included those phases of design that occurred between the end of the feasibility study and the release of detailed drawings to manufacturing.
3) The projects enjoyed good management, i.e., the amount of nonproductive time was small.

---

[5] The terms: effort, design effort, and total design effort are used interchangeably in this thesis.

[6] This has been found to be consistent with practical experience with the average monthly time off due to holidays and vacations (Boehm, 1981).

## 3.2 Factors Affecting Design Effort

Lessons learned in software development indicate that estimation models should include not only product-related factors, but also project-related factors (e.g., see Walston and Felix, 1977; Boehm, 1981; DeMarco, 1982; Jeffery, 1987). These factors should be identified among the more than one hundred factors which influence different aspects of the design process (Hales, 1987; Wallace and Hales, 1987). Nevertheless, after reviewing much previously published research (Walston and Felix, 1977; Boehm, 1981; Cooper, 1990; Griffin, 1993, 1997; Jeffery, 1987; Hajek, 1984; Hales, 1987; Jones, 1986; Wallace and Hales, 1987; Blessing, 1994; Bahill and Chapman, 1995; Waldron and Waldron, 1996; Jacome and Lapinskii, 1997), the factors described below were identified as the most significant ones (Bashir and Thomson, 1999a):

- product complexity
- technical difficulty
- team expertise
- management complexity
- use of automated design tools
- design process.

The above factors are described in detail below. Note, however, that it is not a comprehensive list of all possible factors. One has to recognize that possible factors vary from environment to environment, and that certain factors will be unique for certain environments. For example, as shown in a subsequent section, it was found that for GE, the type of drawings submitted to the customer was a significant factor that affected design effort. Moreover, the number of factors to be included in a model depends on the characteristics of the projects in the data set. However, within one design group, the projects undertaken are often quite similar, and only a few factors need to be considered.

### 3.2.1 Product Complexity

Product complexity, which also reflects project size, is the most significant factor that has an impact on design effort. The relationship between product complexity and effort is obvious. The more complex a product is, the more effort that is required to design it.

Most software effort estimation models include this factor as a dominant parameter. According to Walverton (1974) and Boehm (1981), project size accounted for about 50% of the variation in software project effort.

### 3.2.2 Technical Difficulty

Technical difficulty is another factor that has an impact on design effort. Technical difficulty may be due to the use of new technology, severity of requirements, or a combination of both. Requirements include properties such as quality, reliability, cost, performance, weight, efficiency, and so on.

### 3.2.3 Team Expertise

Team expertise is the main parameter for indicating design team capability. Team (individual) expertise has a direct effect on the effort needed for a project, and more expertise has a positive effect on the efficiency of performing a project. This is because expert individuals handle information more efficiently, spend less time to set the physics of a problem, and generate more solutions than inexpert individuals (Blessing, 1994).

### 3.2.4 Management Complexity

Management complexity has to do with factors that make it more difficult to manage design groups. This has to do with complex reporting and communication structures within the same organization or with design partners. In a study by Hales (1987), it was found that more than 35% of the total design effort was spent in direct communication of some sort or another.

### 3.2.5 Use of Automated Design tools

No doubt that tools such as CAD and other automated design tools have an impact on design effort, especially for large projects. For example, in some applications, the use of CAD has led to an improvement of 300% in productivity compared with manual drafting (Gott, 1980).

### 3.2.6 Design Process

Because the efficiency of a project depends on how well it is carried out, then, design process, which embodies work habits and procedures, is another important factor that has an influence on design effort.

### 3.2.7 The Selected Factors

Inclusion of one or more factors in a model depends on the characteristics of the historical projects in the data set. Based on the data collected from the three companies (Appendix I) along with consultations with their project managers, the following factors were selected to be included in the effort estimation models described in this thesis:

<u>general factors</u>

- product complexity
- severity of requirements
- technical difficulty
- team expertise

<u>company specific factors</u>

- type of drawings submitted to the customer
- involvement of design partners.

The following may be noted with reference to the above selected factors.

- Product complexity was included in all the models described in this thesis.
- Severity of requirements was included in the models constructed for the data collected from NN and CMC.
- The last four factors were included in the models constructed for the data collected from GE.
- Technical difficulty and team expertise were included in the models constructed for the data collected from GE as one combined variable, *difficulty to expertise ratio*.
- The first four factors are general factors that could affect any type of projects, while the last two factors are unique to GE.

Since the accurate estimation of the effort to complete a project requires a realistic appraisal of the complexity (Hajek, 1984), the focus of the following chapter is to develop an objective metric for estimating product complexity. However, to each of the other factors, numerical values were assigned as follows.

**Severity of requirements**[7]

1: design requirements were not too difficult to meet

2: design requirements were difficult to meet

3: design requirements were extremely difficult to meet

**Technical difficulty to expertise ratio**

< 1: if the design was not difficult with respect to the expertise of the team

> 1: if the design was difficult with respect to the expertise of the team

= 1: otherwise

**Type of drawings submitted to the customer**

1: basic drawings

2: assembly drawings

3: manufacturing level drawings

**Involvement of design partners**

1: no design partners were involved

2: design partners were involved

## 3.3 Evaluating a Model

There are a number of objective criteria that can be used for the evaluation of a model. The most widely used criteria include the mean magnitude of relative error (*MMRE)*, prediction at a given level *PRED(l)*, and the coefficient of multiple determination ($R^2$). These criteria are widely used by software researchers (Conte et al., 1986). Because these

[7] For the model described in Chapter 7, numerical values were assigned in a different way.

criteria are often in disagreement and there is no general acceptance of any specific one, all of them were used to evaluate the proposed estimation models.

### 3.3.1 The Mean Magnitude of Relative Error (*MMRE*)

One criterion to test the validity of a model is to examine the mean magnitude of relative error, defined as:

$$MMRE = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|\hat{E}_i - E_i\right|}{E_i} \tag{3.1}$$

where:

$\hat{E}_i$ = *estimated effort of project i in hours or man-months*

$E_i$ = *actual effort of project i in hours or man-months*

$N$ = *number of projects*

A small *MMRE* indicates that on average, the model is a good predictor. According to the Purdue Software Metrics Group (Conte et al., 1986), the model is considered to be acceptable if its *MMRE* is equal to .25 or less.

### 3.3.2 Prediction at a Given Level (*PRED(l)*)

This criterion is used as an indicator of how many of the predicted values fall within a given range of the actual value. In the study, the model is considered to be acceptable if (*PRED*(.25)) ≥ .75. In other words, the model is said to be acceptable, if 75% of the predicted values are within 25% of their actual values (Conte et al., 1986).

### 3.3.3 The Coefficient of Multiple Determination ($R^2$)

This criterion shows the percentage of variance accounted for by the independent variables. A high value of $R^2$ means that a large percentage of variance is accounted for, and additional independent variables are not likely to improve the model much. $R^2$ can be computed by the following formula:

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (\hat{E}_i - E_i)^2}{\sum_{i=1}^{N} (E_i - \overline{E}_i)^2} \tag{3.2}$$

where:

$\hat{E}_i$ = *estimated effort of project i in hours or man-months*

$E_i$ = *actual effort of project i in hours or man-months*

$\overline{E}_i$ = *the mean of the values* $E_i$

### 3.4 Summary

This chapter has described the general methodology adopted to construct the models described in this thesis. This methodology involved data collection, selecting the major factors that influenced design effort, and generation of a model that reflected the relationship among one or more of the selected factors and design effort. Using a number of criteria including the mean magnitude of relative error (*MMRE*), prediction at a given level *PRED*($l$), and the coefficient of multiple determination ($R^2$), models were tested for whether they produced reasonable, accurate estimates. If a model proved accurate, the model was said to be acceptable; otherwise, the formulation of the model was revised.

The data that were used for developing and testing the models described in this thesis were obtained from three companies, namely, Nortel Networks, Canadian Marconi Company, and General Electric Hydro. Based on the collected data along with consultations with project managers, the major factors that should be included in the models were identified. These factors included: product complexity, severity of requirements, technical difficulty, team expertise, type of drawing submitted to the customer, and involvement of design partners. While the first four factors can be considered as general factors that may affect any type of projects, the last two factors are company specific and are unique to certain type of projects, viz., projects that develop engineered-to-order products.

Due to its importance as a dominant parameter in design effort estimation, the focus of the next chapter is on developing a meaningful metric for estimating product complexity. For the other factors, numerical values were assigned. The factors and their corresponding possible values are summarized in Table 3.1.

**Table 3.1** Secondary factors used in the estimation of design effort and their possible values

| Factor | Scale/ value |
|---|---|
| Severity of requirements | 1-3 |
| Technical difficulty to expertise ratio | 1, < 1, > 1 |
| Type of drawing submitted to the customer | 1-3 |
| Involvement of design partners | 1-2 |

# Chapter 4

# A Metric for Estimating Product Complexity

Metrics[8] numerically characterize some attribute of an entity (Fenton and Pfleeger, 1997). Precise and familiar metrics can be seen everywhere in our daily life, e.g., weight, size, temperature, etc. In any discipline, metrics play a vital role. Without metrics, comparisons and predictions are very difficult to achieve. In some disciplines, the development of metrics is not a difficult task; however, it is very difficult in areas of high abstraction such as software development and the general design process. This is because the activity is a mental process without readily identifiable or tangible values. In spite of this difficulty, a significant body of work does exist on the use and benefits of metrics for software development, e.g., Boehm (1981), Jones (1986), Fenton and Pfleeger (1997), and Cote et al. (1988).

Generally, there are two approaches to the development of such metrics, the inductive approach and the deductive approach. The former depends on a considerable amount of observation and/or experimentation; the latter depends on a set of criteria that a metric should satisfy (Elmaghraby and Herroelen, 1980). In this chapter, using the deductive approach, a new metric which estimates functional complexity is proposed to assess product complexity. The new metric has two main uses.

- It allows comparison among design tasks; in other words, complexity can be considered as a design attribute, which is easily estimated.
- Most importantly, it can be used in conjunction with other metrics to estimate the required design effort.

[8] In this thesis, the term metric and measure are considered as synonyms.

## 4.1 Complexity

The term complexity is defined as being proportional to the expected number of man-weeks required to complete tasks (Norden, 1964). Implied by this definition is that, as design complexity increases, the rate of consumption of resources increases. Therefore, if complexity can be measured, then, the required resources can be estimated. Due to the unavailability of acceptable quantitative metrics, most companies do this task subjectively. However, there is a myriad of disadvantages to using subjective estimates. They can be inaccurate, biased and ill suited to sensitivity analysis. (Conte et al., 1986). Because of this, such estimates often lead to unrealistic plans, and thus, project failure (Gioia, 1996). To improve this situation, there is a need to define a good metric to estimate product complexity.

## 4.2 Estimating Product Complexity

Estimating the number of parts to be designed is the simplest way for estimating product complexity. However, there are different theories in the literature concerning the effect of the number of parts on complexity that effect time. For example, some studies indicate that an increase in number of parts leads to an increase in time, and vice versa (Gomory, 1989; Millson et al., 1992; Murmann, 1994). On the other hand, it has been argued that fewer parts can increase the complexity of the remaining parts resulting in an increase in time (Clark 1989, Ulrich et al. 1993). These contradictory conclusions make questionable the use of the number of parts as a complexity metric. Because of this, functionality has received more attention as an alternative and promising aspect to estimate product complexity. This is based on the following.

- The relationship between functionality and resource consumption is obvious. The more functionality that is required, the more complex a product is (Suh, 1990; Griffin, 1993), and thus, more resources are required to design it (Norden, 1964; Griffin, 1993).
- Functionality is based on user requirements; therefore, it is independent of the methodology applied to design a product.
- Functionality has emerged as a very important product attribute in disciplines such as software development, and has been used to good effect in estimating resource

requirements in a number of industrial applications. Function points, conceived by Albrecht (1979), as a metric for effort estimation for software development has proven successful (Fenton and Pfleeger, 1997). This metric estimates the complexity in terms of a weighted sum of delivered functional units. Functional units are defined as the number of inputs, the number of the outputs, the number of inquiries, and the number of files. According to Dreger (1989), the use of function points allows managers to reliably estimate to within 20% of actual time and cost. Furthermore, it was estimated that more than 500 companies rely on this metric. Because the end objective of product development for software and hardware is to deliver functions that satisfy customer needs, the use of a single measure of functionality would be very useful for estimating the amount of effort needed for the development of any product. Unfortunately, the direct use of function points as a metric does not transfer well to domains outside software.

## 4.3 Functionality

The functionality of a product comes from the functions that it delivers to meet design requirements. Design requirements are demands and wishes that clarify the design task in the space of needs (Pahl and Beitz, 1984). Thus, a function can be defined as "the behavior which is required for the device to satisfy a given requirement" (Kota and Ward, 1990). Teleology, design intent, purpose and utility are alternative terms for the intuitive idea of function (Kannapan, 1995).

Two studies have dealt with the issue of measuring product complexity in terms of functionality; one by Griffin (1993), the other by Kannapan (1995). In the former study, the objective was to develop models that could be used to estimate product development time, while design evaluation was the objective of the latter study. Both studies proposed to measure complexity in terms of the number of functions, which are considered as the prime reason for existence of the product. For example, as discussed by Griffin (1993), a vacuum cleaner has two functions, removing dirt and storing dirt; therefore, its complexity is two. From the example by Cross (1994), because it has two functions, removing dirt, and removing excess water, the complexity of a washing machine is two. If another function is added such as drying clothes, then, its complexity is three.

### 4.4 Complexity Metric Criteria

When considering the capabilities of metrics, it is useful to consider them in terms of certain characteristics. Five criteria are given below which describe the characteristics that good metrics for estimating product complexity should have. The characteristics are: intuition, sensitivity, consistency, generality, and simplicity.

**Criterion 1 (intuition)**

A metric should conform to intuition. For example, if a product was considered more complex than another from previous experience, then, the same conclusion should be indicated when a metric is applied.

**Criterion 2 (sensitivity)**

A metric should not be too coarse so as to rate too many products as being of equal complexity, and not be too sensitive so as to assign every product a unique rating.

**Criterion 3 (consistency)**

The complexity of a part must be less than that of the whole. In other words, if x is a component of a product y, then, the complexity of x must be less than that of y.

**Criterion 4 (generality)**

A metric should be applicable to any product. For example, number of integrated circuits does not satisfy this criterion because it is applicable only for electronic products.

**Criterion 5 (simplicity)**

A metric should be simple and easily interpreted.

## 4.5 The Weakness of 'Number of Functions' Metric

The 'number of functions' metric considers only the number of functions to be delivered to the customer. To paraphrase Boehm (1981), this is like "estimating the cost of an automobile by its gross weight or by the number of parts in it". In other words, this metric is not realistic since it is insensitive to the complexity of each function and the relative difficulty of developing functions that are more complex. Therefore, it is clear that this metric does not satisfy criterion 1. Furthermore, the ranking of 36 products as equally complex in Griffin (1993) confirms that this metric is too coarse and does not satisfy criterion 2. In addition, this metric is not consistent. It equates the complexity of the whole and the complexity of the part. For example, in spite of a bulb being part of a pocket flashlight, this metric considers them equally complex (each of them has a complexity of 1). This indicates that the metric does not satisfy criterion 3.

The failure of the 'number of functions' metric to satisfy these three criteria makes questionable its usefulness as an appropriate complexity metric. One approach to make this type of metric more useful is to incorporate other parameters, which would help to rate the relative difficulty of developing different functions. The number of sub-functions and their depth are such parameters. This concept is based on functional decomposition.

## 4.6 Functional Decomposition

The concept of functional decomposition was invented by Larry Miles during World War II (Miles, 1961). Since then, it has had a long history of application as a tool for analyzing the performance and usefulness of a product or service (Shillito and De Marle, 1992). Since the 1980's, there has been a rising emphasis on the use of this concept in the design process (Pahl and Beitz 1988; Kota and Ward, 1990; Suh 1990; Hubka, 1988; Kusiak and Szczerbicki, 1992; O'Shaughnessy and Sturges, 1992). In the functional decomposition process, each function that the product to be designed must perform is decomposed into sub-functions. Then, each sub-function is further broken down into sub-functions, and so on.

The relationship between functions and sub-functions can be represented in different forms such as a block schematic or a functional tree. However, because of its simplicity, a functional tree is most widely used (Hubka, 1988). Any functional tree consists of blocks that are connected by branches. Each block represents a basic function if it is at the first or highest level or a sub-function if it is at a level lower than the first. The number of levels in the decomposition is indicative of the complexity of the design task (Hubka, 1988; Kota and Ward, 1990; Kusiak and Szszerbicki, 1992). For example, because its functions cannot be decomposed into further levels of sub-functions, a product such as a bolt is considered to be one of the least complex products. On the other hand, since its functions can be decomposed into a large number of levels of sub-functions, a product such as an electricity power plant is considered to be one of the most complex products. This property can be used to advantage in determining how to measure product complexity.

## 4.7 The Proposed Complexity Metric

If it is assumed that product complexity depends on the number of functions and the depth of their functional trees (hierarchies), then, a metric, *PC*, for product complexity can be defined by the following formula (Bashir and Thomson, 1999b):

$$PC = \sum_{j=1}^{l} F_j j \tag{4.1}$$

where:

$F_j$ = *number of functions at level j*

$l$ = *number of levels*

### 4.7.1 General Guidelines

For the product complexity metric, *PC*, to be effective, functional decomposition needs to be performed consistently. Before the method of decomposing product functions is presented, it is helpful to outline the following general guidelines. Unless indicated, function is used as a general term to indicate a basic function or a sub-function.

- Any function must be expressed through a verbal model that combines one verb and one noun (Pahl and Beitz, 1984), for example, increase temperature, increase speed, or hold material. If a function cannot be described as one verb and one noun, this likely indicates that more than one function exists or it is not a function at all (Bronikowski, 1986).
- It is possible to have a function that can be achieved in different ways. Therefore, it is possible to have different design solutions (alternatives), each of which is represented by a different functional tree. The functions in the lower level must not be decomposed unless all the functions at the higher level have been considered.
- If a given function cannot be further decomposed into simpler functions (verbal models), or if it can be matched with an existing component without any change, or if a component will be designed by a subcontractor, then no further decomposition is required and there will be no further lower levels for that function. Thus, the depth of decomposition indicates the complexity of the function or lack of existing physical components to fulfill the function. The depth of decomposition is therefore an index of the degree of product newness.

### 4.7.2 Decomposition Steps

As shown in Figure 4.1, the decomposition steps are summarized as follows:

1) From design requirements, the overall basic functions (first-level functions) are determined and placed at the highest level of the functional tree. The basic functions are those that must be performed by the product.
2) Once all the basic functions have been identified, they are decomposed one by one into sub-functions. The decomposition is achieved by determining all the functions that must be done to accomplish the corresponding basic function. The sub-functions are placed at the next level down.
3) Once all the basic functions are considered, the sub-functions at the next level down (second-level functions) are further decomposed one by one into sub-functions similar to what was done in the previous level, and so on.
4) Completing the above steps will result in a functional tree that contains all the sub-functions needed to accomplish the corresponding basic functions. Since the

relationships between the basic functions as well as the sub-functions are AND/OR type, then logically, it is possible to extract from each tree a number of alternatives. Then, using the formula for product complexity, *PC*, the complexity of each design alternative can be assessed.

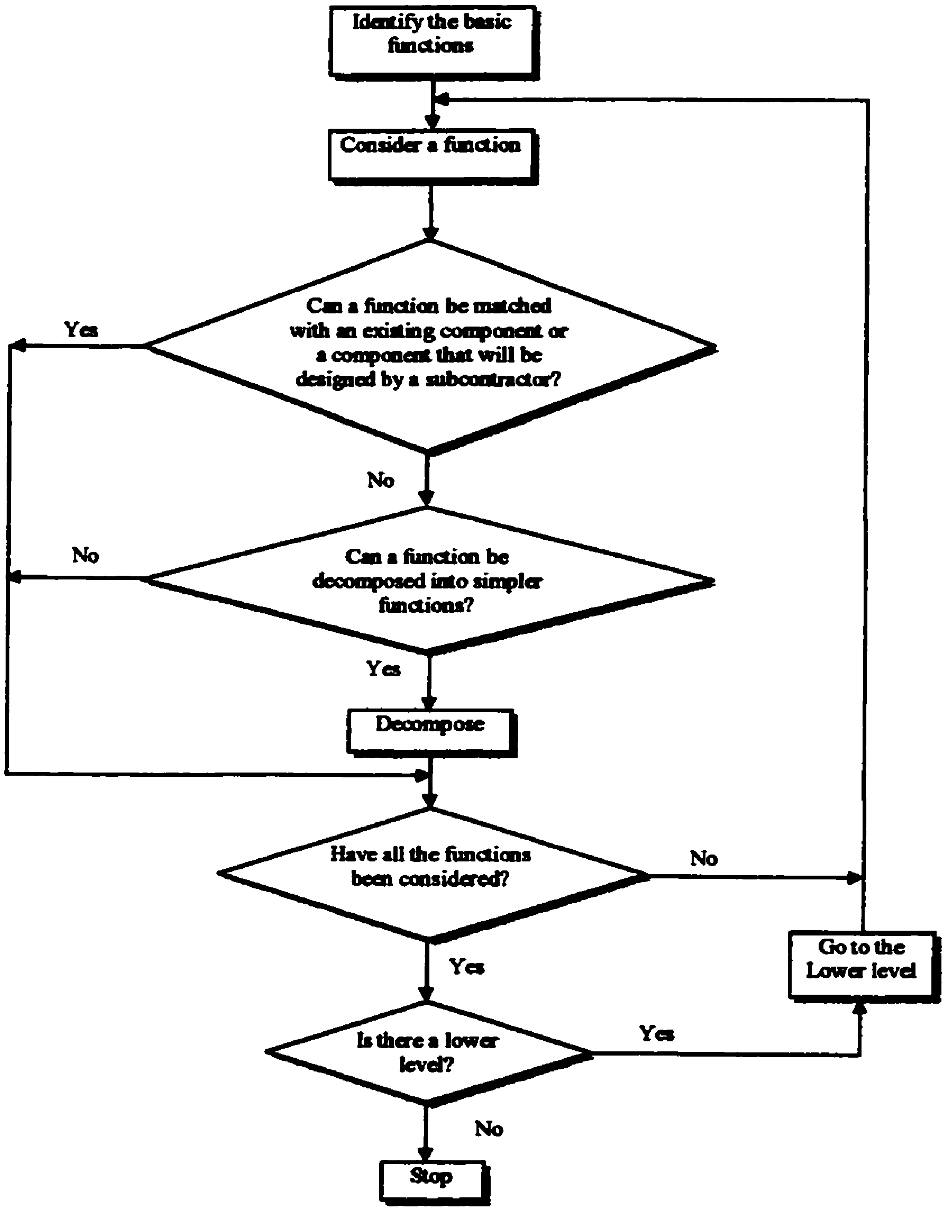



**Figure 4.1** Functional decomposition steps

### 4.7.3 Alternative Method for Estimating Product Complexity

In addition to the general guidelines given in Section 4.7.1, if a weight can be confidently assigned to differentiate between the complexity between a given function at a certain level of decomposition and the complexities of the corresponding functions in previously designed products, then, no further decomposition for that function is required. In this case, *PC*, is computed using equation (4.2).

$$PC = \sum_{j=1}^{l} f_j j \qquad (4.2)$$

where:

$f_j = \sum_{k=1}^{F_j} w_k$

$w_k$ = *weight assigned to function k*

$F_j$ = *number of functions at level j*

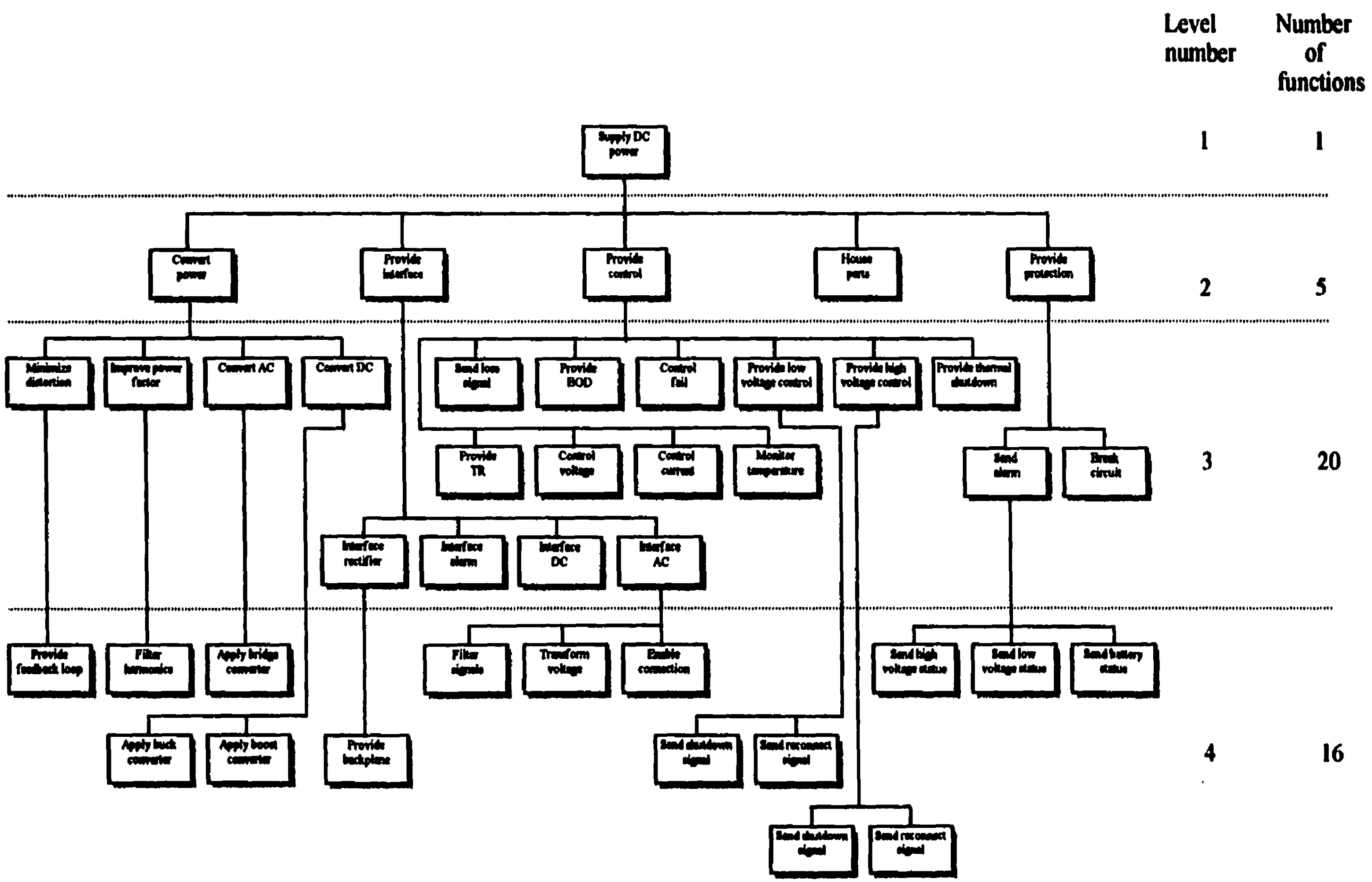
$l$ = *number of levels*

If each function is given a weight of 1, then equation (4.2) reduces to equation (4.1).

### 4.7.4 Demonstrative Examples

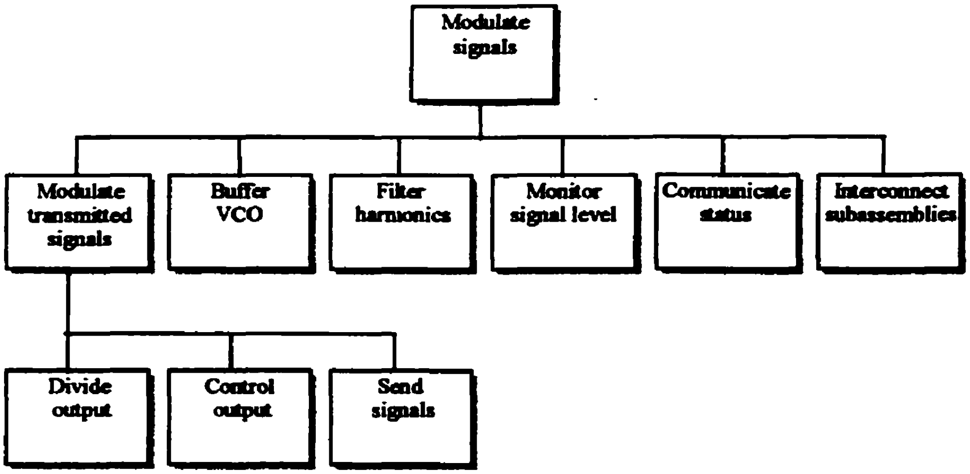
To test their applicability, the above guidelines and steps were followed by designers from the three companies to construct the functional trees of a sample of previously designed products. The constructed functional trees and their computed complexities are shown in Figures 4.2-4.5. Note that:

- The functional tree in Figure 4.2 corresponds to a battery charger, while the functional trees in Figures 4.3 and 4.4 correspond to a modulator and a radio frequency unit, respectively, where the former is a part from the latter.
- The functional tree in Figure 4.5 corresponds to a generator where weights are given to certain functions to differentiate their complexities from corresponding functions in other designed products.




$$PC = (1 \times 1) + (5 \times 2) + (20 \times 3) + (16 \times 4) = 135$$

**Figure 4.2** The functional tree of a battery charger and its computed complexity




$PC = 22$

**Figure 4.3** The functional tree of a modulator and its computed complexity

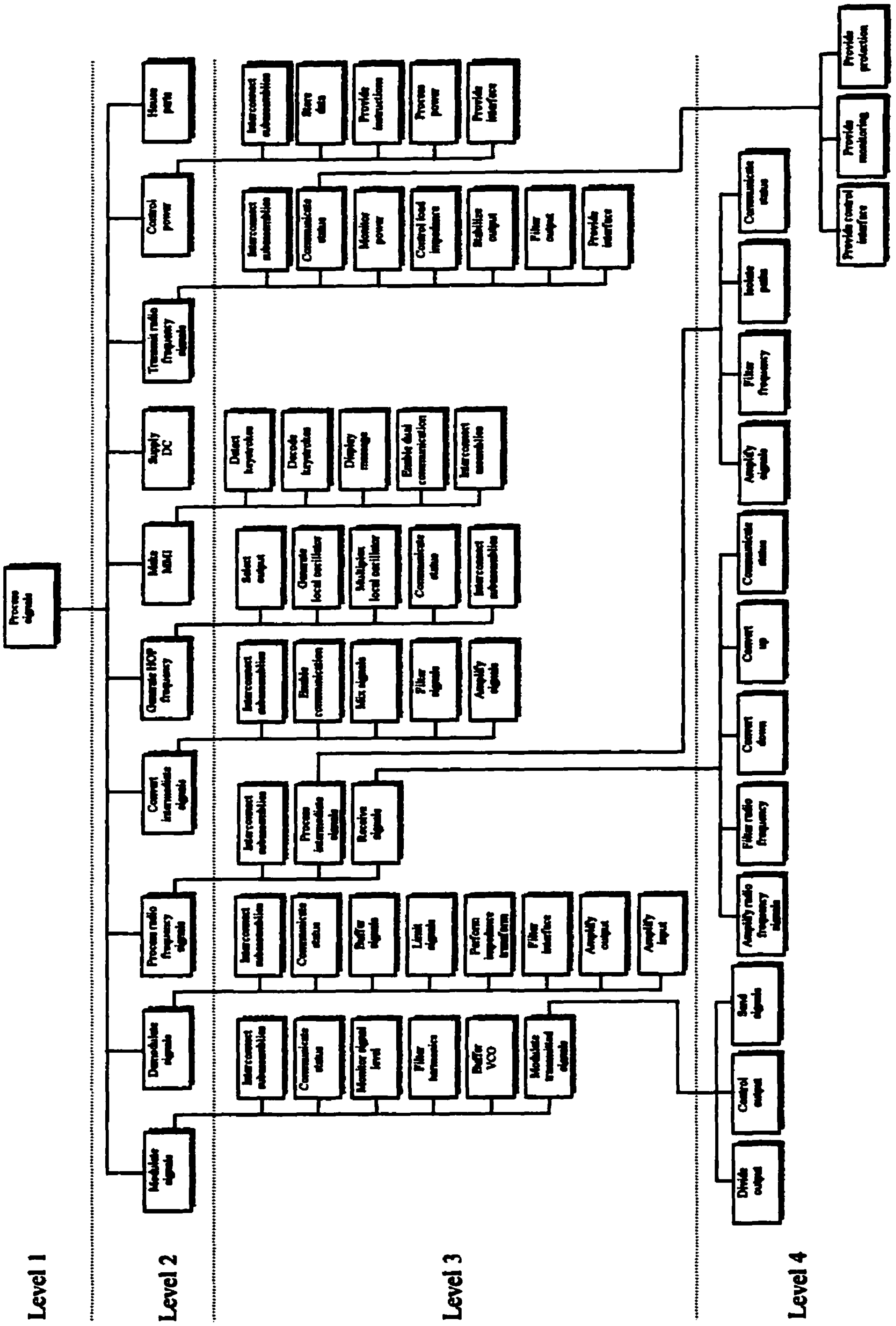
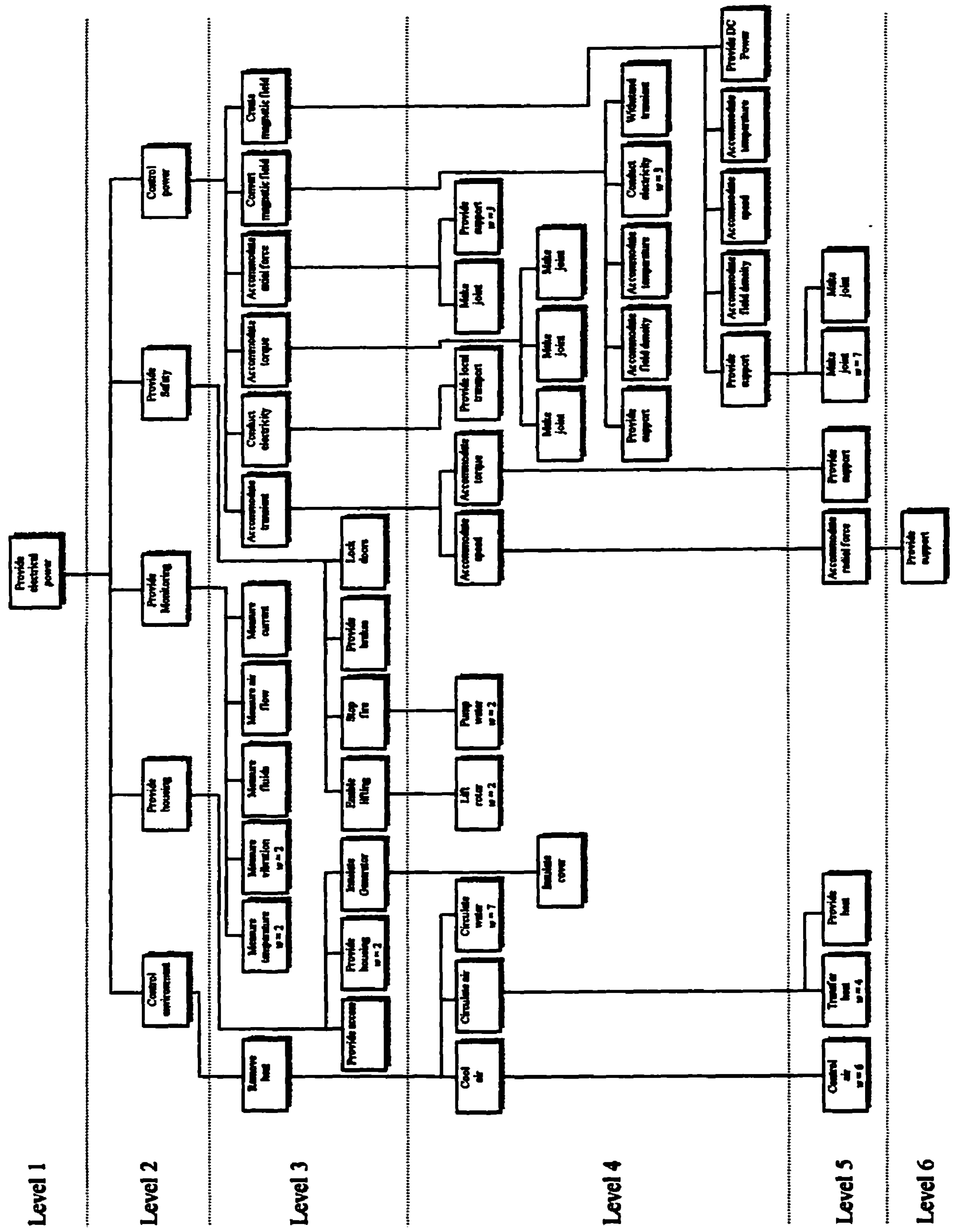



**Figure 4.4 The functional tree of a radio frequency unit and its computed complexity**




$PC = (1 \times 1) + (5 \times 1) \times 2 + (16 \times 1) + (3 \times 2) \times 3 + \ldots + (1 \times 1) \times 6 = 332$

**Figure 4.5 The functional tree of a generator and its computed complexity**

These figures clearly indicate that the metric is somewhat intuitive. The more sub-functions overall and the more at any level, the greater the complexity. The greater the depth of the functional tree, i.e., the more levels, the greater the complexity. Thus, the metric counts the number of functions at each level and weights them by the number of the level.

According to the given guidelines, each of the lowest level functions in Figures 4.2-4.4 meets one or a combination of the following conditions:

- it was mapped to a component which was designed by a subcontractor,
- it was mapped to an existing component, and/or
- it was considered simple.

While each of the lowest level functions in Figure 4.5 meets the following conditions:

- one or a combination of the above three conditions, or
- it was assigned a weight. For example, there are many ways to 'control air', the function at level 5. It was simpler to assign a weight than decomposing the function further for different designs.

It is worth mentioning that since each of the products delivers one function to the customer, the 'number of functions' metric rates the products shown in Figures 4.2-4.5 as equally complex[9]. Nevertheless, the complexities of these products are different, and this is captured by the proposed metric.

Furthermore, since the complexity of a function at any level is a function of the number of functions at the lower levels and their number of levels, the metric is considered to be consistent. In other words, the metric always indicates that the complexity of a part is always less than that of the whole. For example, since the product in Figure 4.3 is a part of that in Figure 4.4, the metric indicates that the former is less complex than the latter.

[9] Even if the number of functions at level 2 of the functional diagrams is used, this value is not enough to differentiate between different complexities.

These observations confirm that the shortcomings of the 'number of functions' metric are overcome by the proposed metric.

## 4.8 Validating the Proposed Metric as a Predictor of Design Effort

To test whether a metric can measure what it claims it can measure, one or a combination of the following approaches are usually adopted: the experimental and the empirical. The former approach uses data from experiments, while the latter approach uses actual data from large-scale projects[10]. In this research, both approaches were adopted to validate the proposed metric as predictor of design effort.

### 4.8.1 Experimental Validation

Conducting an experiment involving design tasks is not easy. This is because design involves much mental activity. Even if the subjects follow clear instructions, it is still an open question as to how closely they will follow them. Furthermore, requiring all subjects to do the same task does not ensure that they will produce the same output. In addition, adopting different design methodologies can lead to substantially different amounts of effort. With these facts taken into account, the following sections describe an experiment which was conducted in the Department of Mechanical Engineering at McGill University to test the causal link between product complexity, *PC*, as estimated by the proposed metric and design effort, *E*.

#### 4.8.1.1 Independent and Dependent Variables

Three levels of design complexity were selected as an independent variable. The time in minutes taken to complete the task was the dependent variable.

#### 4.8.1.2 Design Tasks

The tasks referred to as task A, task B, and task C involve designing simple devices for positioning a workpiece at a desired position for welding operations[11]. The descriptions of

[10] Projects where many people are involved and which take months or years, instead of hours or days, to complete.

[11] They are modified versions of a case study presented in Hubka (1982).

these tasks are presented in Appendix II. The estimated product complexities, *PC*, of the positioning devices were 11, 22, and 33, respectively.

#### 4.8.1.3 Subjects

Twenty eight subjects participated. To eliminate the effect of the variance due to the possible differences in their design abilities, all the subjects were graduate and upper-level undergraduate mechanical engineering students at McGill University. The subjects were paid for performing the tasks.

#### 4.8.1.4 Design of the Experiment

The design of this experiment is called the one way model, since each subject is given a single task. The subjects were assigned at random to the different tasks. Often this type of experimental design is referred to as completely randomized design[12]. One shortcoming of this design is that even though subjects are randomly assigned to the tasks, it is possible to assign more experienced subjects to one condition than to another quite by random (Weimer, 1995). However, in this experiment it was assumed that there were no significant differences in the expertise of the subjects.

#### 4.8.1.5 Procedure

The subjects were provided with all necessary references and tools to perform the assigned tasks. Before starting the experiment, they were given written and oral instructions on how to perform the tasks (see Appendix II). In addition, they were given enough time to train themselves on how to use the provided references. To eliminate the effects of fatigue, the subjects were allowed to take breaks whenever they began to feel tired. Break time was excluded from the measured time of effort. There was no time limit for completing the assigned tasks. In other words, the students themselves decided when the assignment was completed.

[12] It is recognized that randomized block design is more effective than completely randomized design. However, completely randomized design was a better choice for this case where carry-over effects may occur.

#### 4.8.1.6 Analysis of the Experiment Results

A professional designer assessed all the design solutions. Meeting the requirements was the assessment criterion. The quality of the drawings was not included as part of the assessment criteria. For the analysis, 8 solutions for task A, 5 solutions for task B, and 5 solutions for task C were selected as achieving relative scores of equal or greater than 60%, i.e., the solutions achieved 60% of the requirements. Scores are shown in Figure 4.6. Time in minutes spent on each of these solutions is shown in Table 4.1. Figure 4.7 represents a scatter diagram for the data in Table 4.1. From examining this scatter diagram, it is clear that the three levels of complexity have different influence on design effort. This visual assessment was supported by a significance testing technique.

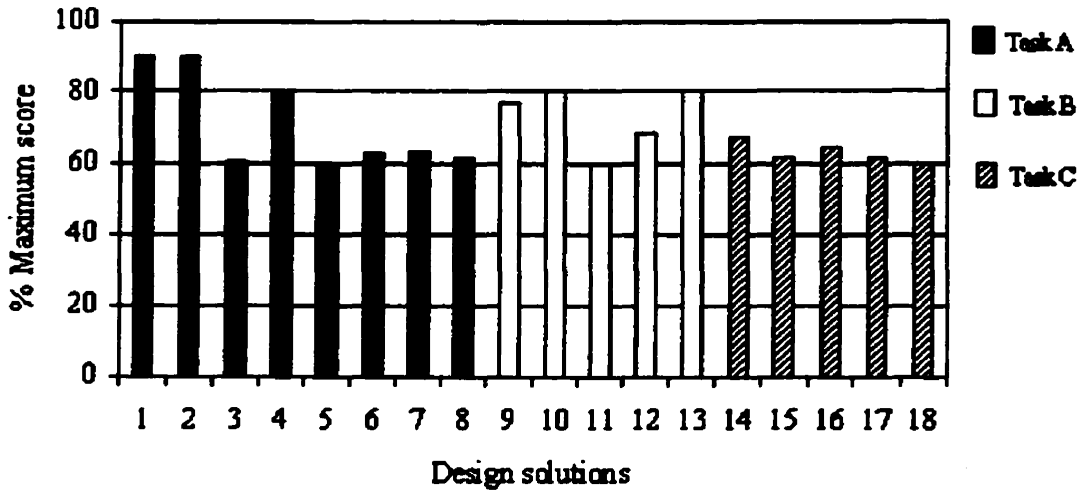



**Figure 4.6** Assessment of the design solutions

**Table 4.1** Design effort, $E$, (in minutes) spent by the subjects

| Task A | Task B | Task C |
|---|---|---|
| 102 | 158 | 181 |
| 112 | 170 | 217 |
| 99 | 150 | 200 |
| 101 | 150 | 189 |
| 150 | 133 | 200 |
| 123 | | |
| 107 | | |
| 99 | | |

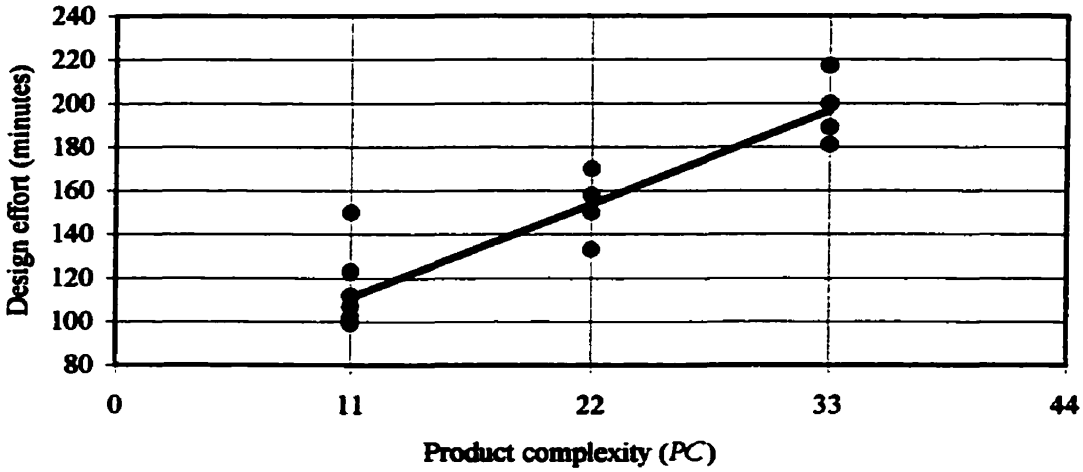



**Figure 4.7** Product complexity versus design effort (experimental results)

Because the distribution of the data is unknown, a distribution-free test known as the Kruskal-Wallis test was adopted for testing the research hypothesis that the three levels of complexity would differ according to their influence on design effort.

#### 4.8.1.6.1 The Kruskal -Wallis Test

The Kruskal-Wallis test is a rank test technique, which can be used in situations where the normality assumption is unjustified (Montgomery, 1984). The technique tests the null hypothesis ($H_o$) that the treatments (levels) are identical against the alternative hypothesis ($H_1$) that some of the treatments are different. In this technique all the observations are ranked in ascending order with the average rank given to each value in a tie; then, the test statistic (*KW*) is computed using equation (4.3).

$$KW = \frac{12}{N_T(N_T+1)} \sum_{j=1}^{k} n_j \overline{R}_j^2 - 3(N_T+1) \tag{4.3}$$

where:

*k* = *number of treatments*

$n_j$ = *number of observations in the $j^{th}$ treatment*

$N_T$ = *total number of observations*

$R_j$ = *average of the ranks in the* $j^{th}$ *treatment*

If there are ties in the observations, *KW* is computed by equation (4.3) and then divided by:

$$1-\frac{\sum_{i=1}^{g} t_i^3 - t_i}{N_T^3 - N_T}$$

where:

g = *number of groupings of different tied ranks*

$t_i$ = *number of tied ranks in the* $i^{th}$ *grouping*

**4.8.1.6.2 Analysis of Results**

Since the hypothesis was that the three levels of complexity would differ according to their influence on design effort, the following is the result of the tested hypothesis by the Kruskal-Wallis at 0.01 level of significance:

$H_o$: there is no difference between the levels of complexity with respect to their influence on design effort

$H_1$: the levels of complexity differ with respect to their influence on design effort

The computed value of *KW* is 14.04. Since $KW > \chi^2_{.01,2} = 9.21$ (see Appendix Table C of Siegel and Castellan, 1988) the null hypothesis, $H_o$, was rejected. Thus, there is a statistically significant difference between the levels of complexity with respect to their influence on design effort. This confirms that there is a causal link between product complexity and design effort.

### 4.8.2 Empirical Validation

An empirical validation was carried out by testing the correlation between product complexity and design effort. The required data are extracted from Appendix I and shown in Tables 4.2, 4.3, and 4.4. As shown in Table 4.2, for NN, the complexity of the designed products, *PC*, ranged from 43 to 135. The effort, *E*, required to design them ranged from 4616 to 25,033 hours. As shown in Table 4.3, for CMC, the complexity of the designed products, *PC*, ranged from 5 to 34. The effort, *E*, required to design them ranged from 632 to 9828 hours. As shown in Table 4.4 for GE, the complexity of the designed products, *PC*, ranged from 308 to 383. The effort, *E*, required to design them ranged from 8192 to 30400 hours.

**Table 4.2** Design effort and corresponding product complexity for a number of projects for NN

| Project number | Design effort (hours) | Product complexity |
|---|---|---|
| 1 | 4616 | 43 |
| 2 | 8800 | 73 |
| 3 | 7500 | 76 |
| 4 | 11468 | 90 |
| 5 | 25033 | 135 |

**Table 4.3** Design effort and corresponding product complexity for a number of projects for CMC

| Project number | Design effort (hours) | Product complexity | Project number | Design effort (hours) | Product complexity |
|---|---|---|---|---|---|
| 1 | 951 | 5 | 7 | 1985 | 11 |
| 2 | 632 | 7 | 8 | 1777 | 15 |
| 3 | 1103 | 7 | 9 | 4950 | 19 |
| 4 | 1099 | 11 | 10 | 3701 | 22 |
| 5 | 1367 | 11 | 11 | 8883 | 24 |
| 6 | 1874 | 11 | 12 | 9828 | 34 |

**Table 4.4** Design effort and corresponding product complexity for a number of projects for GE

| Project number | Design effort (hours) | Product complexity | Project number | Design effort (hours) | Product complexity |
|---|---|---|---|---|---|
| 1 | 20392 | 383 | 9 | 19824 | 336 |
| 2 | 8192 | 322 | 10 | 16944 | 332 |
| 3 | 13544 | 322 | 11 | 20112 | 350 |
| 4 | 11880 | 308 | 12 | 26816 | 362 |
| 5 | 8384 | 319 | 13 | 10704 | 326 |
| 6 | 27200 | 368 | 14 | 10856 | 335 |
| 7 | 20800 | 332 | 15 | 8760 | 279 |
| 8 | 30400 | 352 | | | |

The plots of actual effort versus product complexity for NN, CMC, and GE are shown in Figures 4.8, 4.9, and 4.10, respectively. These graphs indicate that there is a positive relationship between product complexity and design effort. To test how significant the relationship was, the Spearman rank-order correlation coefficient, $r_s$, was computed for the data from the companies. The Spearman rank-order correlation coefficient is a robust measure of association that can be used with data, which are not normally distributed. In addition, it has the advantage of being resilient both to atypical values and to non-linearity of the underlying relationship. To use this measure, ranks are obtained by putting the attribute values into ascending order and giving the smallest value the rank of value 1, the next rank value 2, and so on. If two or more values are equal, they are given the average of the related rank values. Then, $r_s$, is computed using equation (4.4). For more details, refer to Siegel and Castellan (1988). The computed $r_s$ for Nortel, CMC, and GE were 0.90, 0.91, and 0.92, respectively. The coefficients were at $p < 0.05$, $p < 0.01$, and $p < 0.01$ levels of significance, respectively. This result confirms that the proposed metric in any form (equation (4.1) or (4.2)) is a good predictor of design effort.

$$r_s = 1 - 6\sum_{i=1}^{n} \frac{dr_i^2}{n^3 - n} \qquad (4.4)$$

where:

*n = sample size*

*$dr_i$ = the difference in ranks of the $i^{th}$ pair of data*

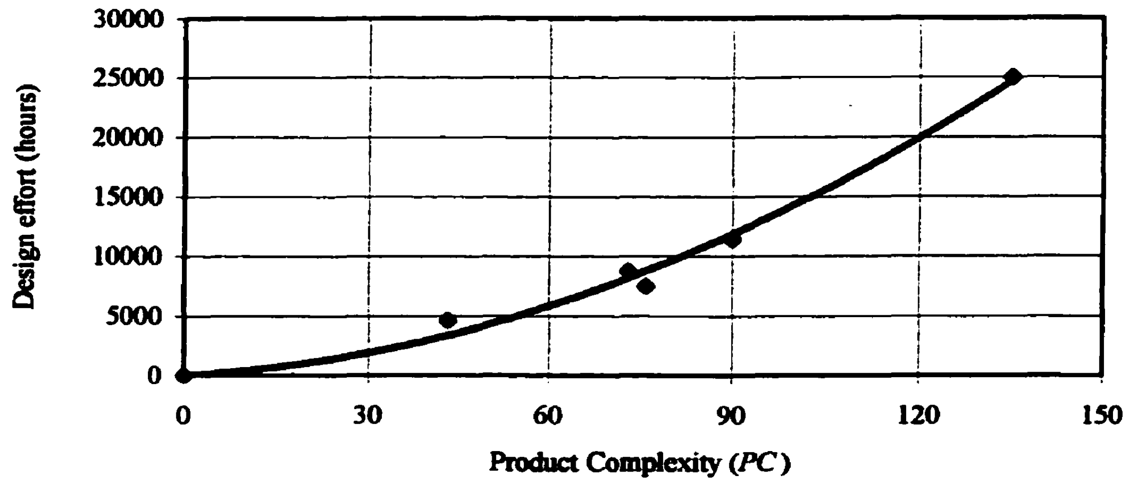



**Figure 4.8** A graph of design effort for projects with different product complexities for NN

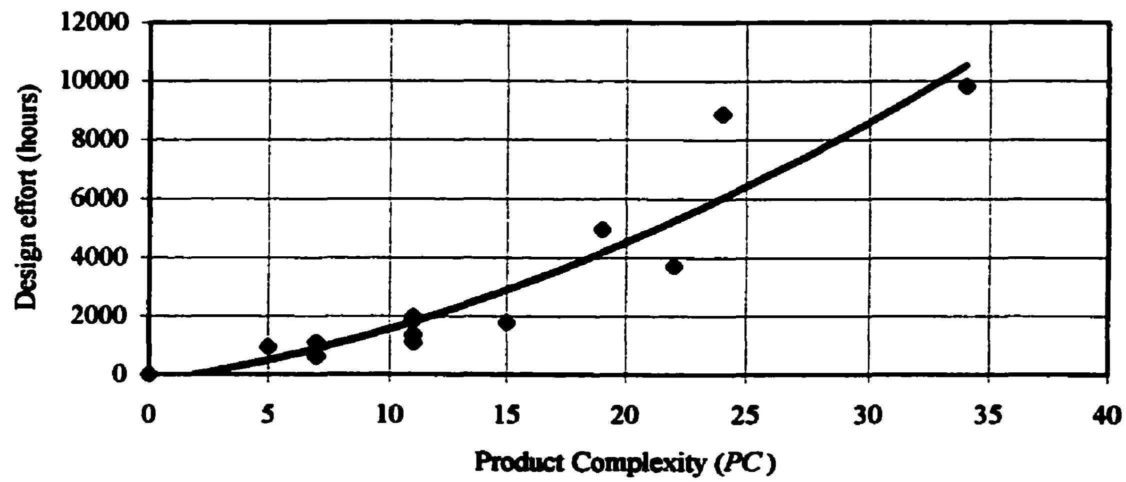



**Figure 4.9** A graph of design effort for projects with different product complexities for CMC

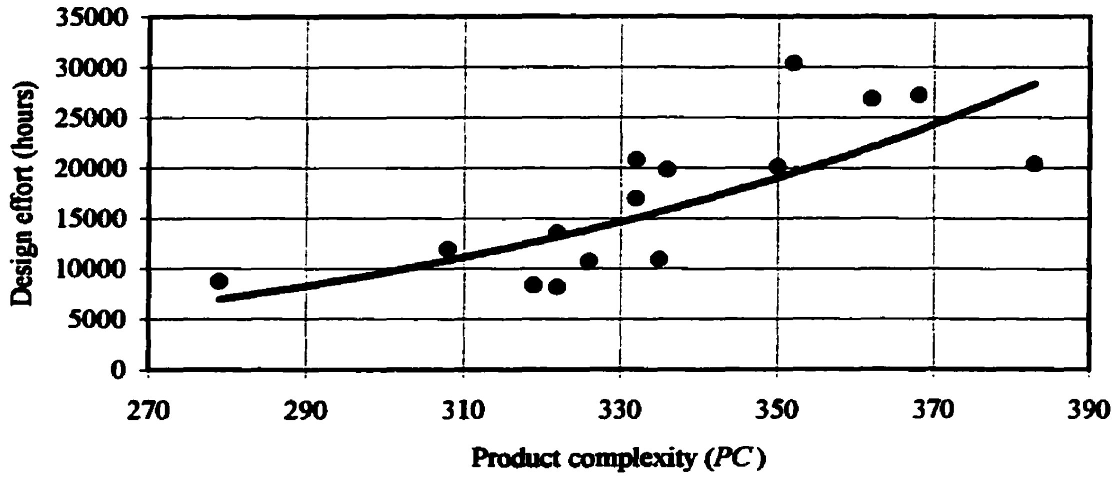



**Figure 4.10** A graph of design effort for projects with different product complexities for GE

## 4.9 Summary

The first step towards developing an effort estimation model is to define a metric that realistically reflects the complexity of the product to be designed. In this chapter, several criteria that should be satisfied by an acceptable complexity metric were introduced. On the basis of these criteria, the shortcomings of the 'number of functions' metric were highlighted.

A new complexity metric, which is based on measuring the complexity of the functional tree of a product using simple weighting factors, was presented.

Based on the proposed protocol for functional decomposition, the size of the functional tree depends on the amount of design to be carried out, the degree of innovation (Kota and Ward, 1990), and the simplicity of the functions.

To validate the proposed metric as a predictor of design effort, both experimental and empirical approaches were adopted. In the former approach, data obtained from a controlled experiment where a number of subjects performed simple design tasks were analyzed using the Kruskal-Wallis test. While in the latter approach, the correlation between product complexity and design effort was tested by the Spearman rank-order correlation coefficient using data collected for a number of projects from NN, CMC, and GE. The results of both approaches confirmed the validity of the product complexity metric, *PC*, as a good indictor of the amount of design effort.

It is worth mentioning that different mathematical forms for product complexity, *PC*, using the characteristics of a functional tree were tried. The forms given by equations (4.1) and (4.2) proved to be the best predictors of design effort.

# Chapter 5

# Parametric Models

Parametric estimation models use historical data from previous projects to establish mathematical relationships capable of generating effort estimates for future projects. In this chapter, using the methodology described in Chapter 3, two types of parametric estimation models were constructed and investigated: single variable models and multivariable models. All the models are based on the product complexity metric described in the previous chapter. The models were developed for the data collected from NN, CMC, and GE using the traditional regression analysis technique.

## 5.1 The Form of Equations

A major step involved in this technique is determining the form of equations. The following general form of equation was selected (Bashir and Thomson, 1999c). It was chosen for simplicity, and the empirical evidence supporting this form (Walston and Felix, 1977; Boehm, 1981; Jeffery, 1987).

$$\hat{E} = a\,PC^{b} D_1^{c_1} D_2^{c_2} \ldots D_m^{c_m} \tag{5.1}$$

where:

$\hat{E}$ = *estimated design effort in hours*

$PC$ = *product complexity*

$D_m$ = *effort driver (factor m)*

$a$, $b$, $c_n$ = *constants (weights) that are estimated from historical data*

Because of the small size of the data samples, the jackknife technique was used to obtain estimates of the equation parameters.

## 5.2 The Jackknife Technique

The jackknife technique is a statistical method, which can be used to ameliorate not only the problem of biased estimates due to the small size of a sample, but also in situations where the distribution for the observations is hard to assess (Eyman et al., 1973; Mosteller and Tukey, 1977). In this technique, the desired calculation for all the data is made, the data are divided into subsamples, and then, the calculation is made for each group of data obtained by leaving out one subsample. Pseudo-values, $Ps_i$, are then calculated using equation (5.2).

$$Ps_i = ns\,\hat{\beta} - (ns-1)\hat{\beta}_{-i} \quad (i = 1,\cdots,ns) \tag{5.2}$$

where:

$Ps_i$ = *pseudo-value for the entire sample omitting subsample i*

$ns$ = *number of subsamples*

$\hat{\beta}$ = *the least squares estimator of the entire sample*

$\hat{\beta}_{-i}$ = *the least squares estimator of the entire sample omitting subsample i*

then, the jackknife estimator, $\tilde{\beta}$, is given by equation (5.3).

$$\tilde{\beta} = \frac{\sum_{i=1}^{ns} Ps_i}{ns} \tag{5.3}$$

## 5.3 The Company Specific Models

As shown below, depending on the number of predictors used, two types of parametric models were constructed and investigated: a single variable model and a multivariable model. A single variable model uses one factor, namely product complexity, *PC*, as a predictor of effort, while multivariable models use one or more of the factors described in Chapter 3 in addition to product complexity.

<u>NN</u>

$$\hat{E} = 24.33PC^{1.39} \qquad (5.4)$$

<u>CMC</u>

$$\hat{E} = 60.44PC^{1.41} \qquad (5.5)$$

$$\hat{E} = 70.65PC^{1.15}\,SR^{0.87} \qquad (5.6)$$

<u>General models using data from NN and CMC</u>

$$\hat{E} = 158.04\,PC^{0.99} \qquad (5.7)$$

$$\hat{E} = 151.38\,PC^{0.82}\,SR^{0.84} \qquad (5.8)$$

<u>GE</u>

$$\hat{E} = 3.2\times10^{-8}\,PC^{4.63} \qquad (5.9)$$

$$\hat{E}_e = 5.9\times10^{-15}\,PC^{7.07} \qquad (5.10)$$

$$\hat{E} = 0.12\,PC^{2}\,DE^{0.41}\,TD^{0.35}\,DP^{0.69} \qquad (5.11)$$

$$\hat{E}_e = 4.8\times10^{-8}\,PC^{4.29}\,DE^{0.9}\,TD^{0.45}\,DP^{0.68} \qquad (5.12)$$

where:

$\hat{E}$ = *estimated design effort in hours*

$\hat{E}_e$ = *estimated engineering effort in hours*

$PC$ = *product complexity*

$SR$ = *severity of requirements*

$DE$ = *difficulty to expertise ratio*

$TD$ = *type of drawings submitted to the customer*

$DP$ = *involvement of design partners*

The following may be noted with reference to the above models.

- Because of the small size (five data points), only a single variable model was developed for Nortel.

- To test the use of one general model to estimate the effort for different companies, equations (5.7) and (5.8) were developed from the combined data from NN and CMC.
- For GE, models that estimate engineering effort, $\hat{E}_e$, were also developed. Engineering effort, $E_e$, comprises the amount of time spent on all design activities excluding drafting[13].

## 5.4 Performance Evaluation Results

Plots of the actual project hours versus hours estimated by the models are shown in Figures 5.1-5.9. Note that the plot for a perfect estimation model would have all the data points on the solid line, which connects the points representing the actual effort. The scatter plots give a quick idea about how the models differ from each other in terms of their performance. A visual assessment of these plots indicates that:

- the multivariable models performed better than their corresponding single variable models, e.g., see Figure 5.2 versus Figure 5.3, and
- the models which were based on fairly homogeneous projects (from one environment) performed better than those which were based on data combined from different environments, e.g., see Figure 5.3 versus Figure 5.5.

However, this is just an informal analysis and may be too subjective to be useful. The best way to evaluate the performance of a model is to use the objective criteria described in Chapter 3, namely, the mean magnitude of relative error (*MMRE)*, prediction at a given level (*PRED*($l$)), and the coefficient of multiple determination ($R^2$). Using these criteria, evaluation of the performance of the models defined by equations (5.4)-(5.12), is discussed in the following section.

[13] Drafting effort can be estimated by subtracting the estimated engineering effort, $\hat{E}_e$, from the estimated design effort, $\hat{E}$.

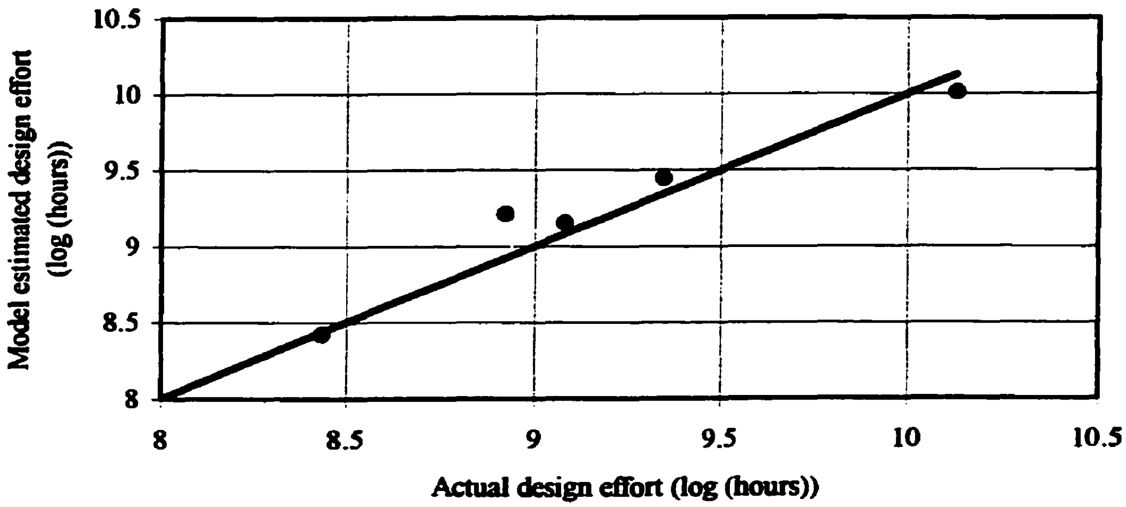



**Figure 5.1** Actual design effort versus single variable model (equation (5.4)) estimates for NN

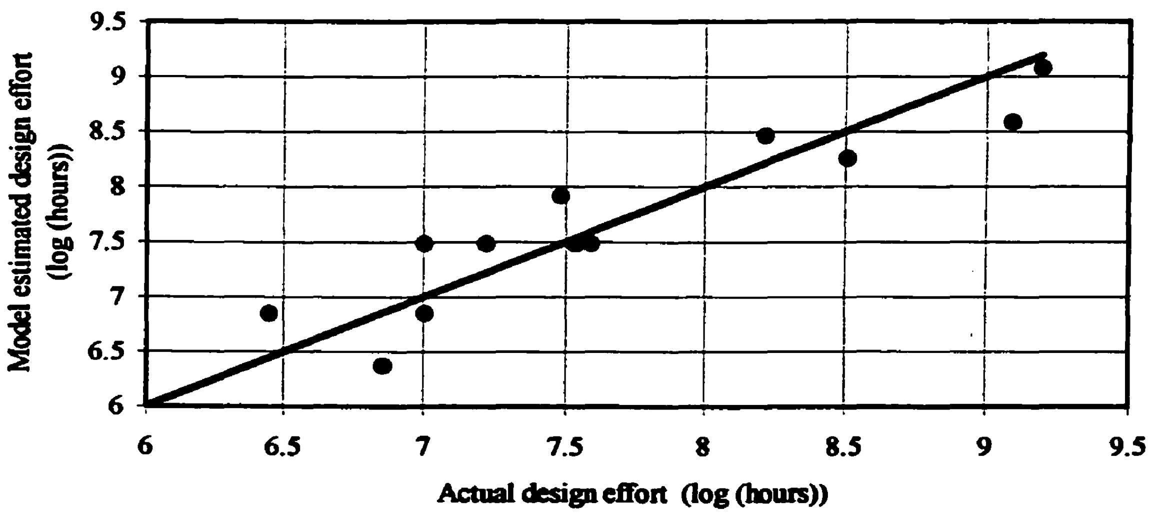



**Figure 5.2** Actual design effort versus single variable model (equation (5.5)) estimates for CMC

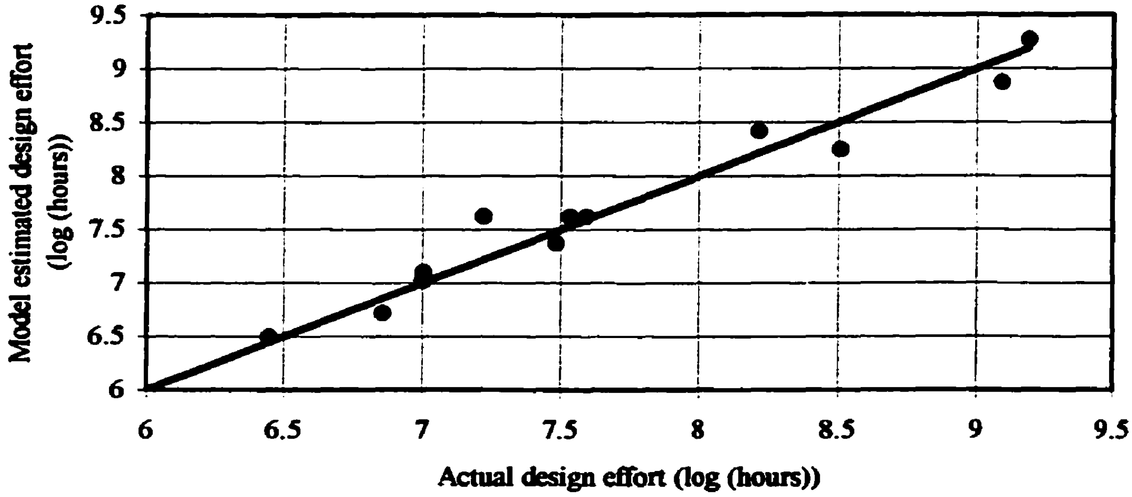



**Figure 5.3** Actual design effort versus multivariable variable model (equation (5.6)) estimates for CMC

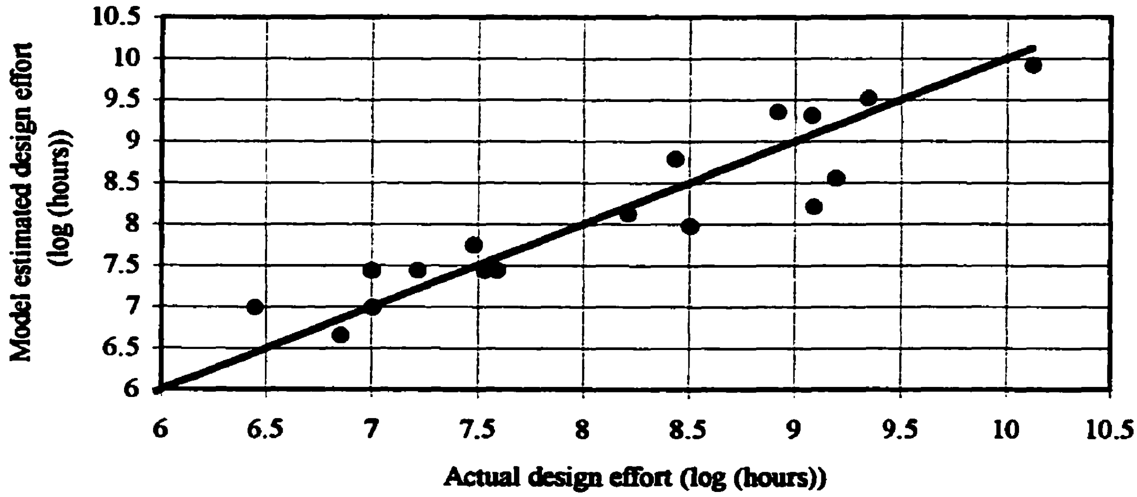



**Figure 5.4** Actual design effort versus single variable model (equation (5.7)) estimates for NN and CMC combined data

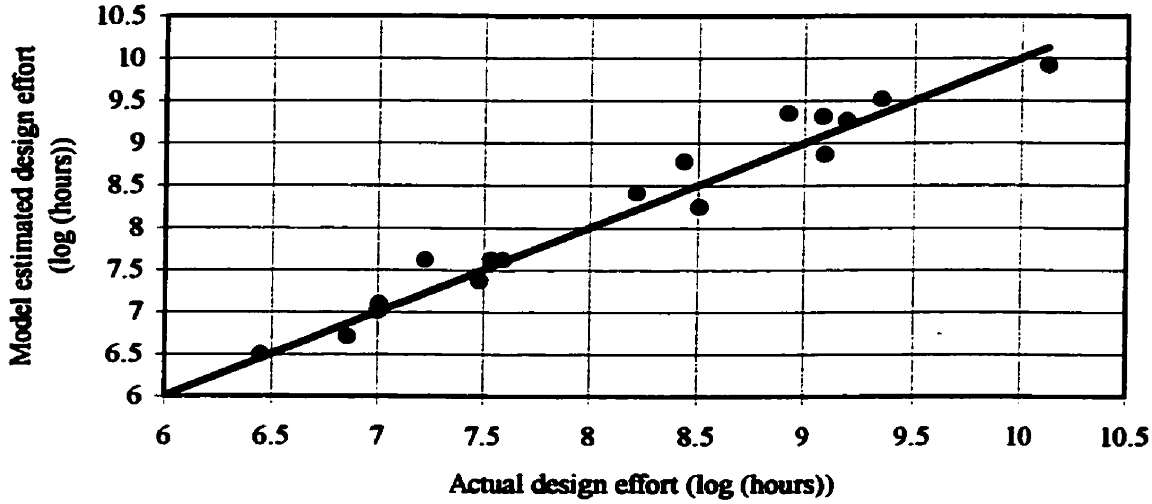



**Figure 5.5** Actual design effort versus multivariable variable model (equation 5.8) estimates for NN and CMC combined data

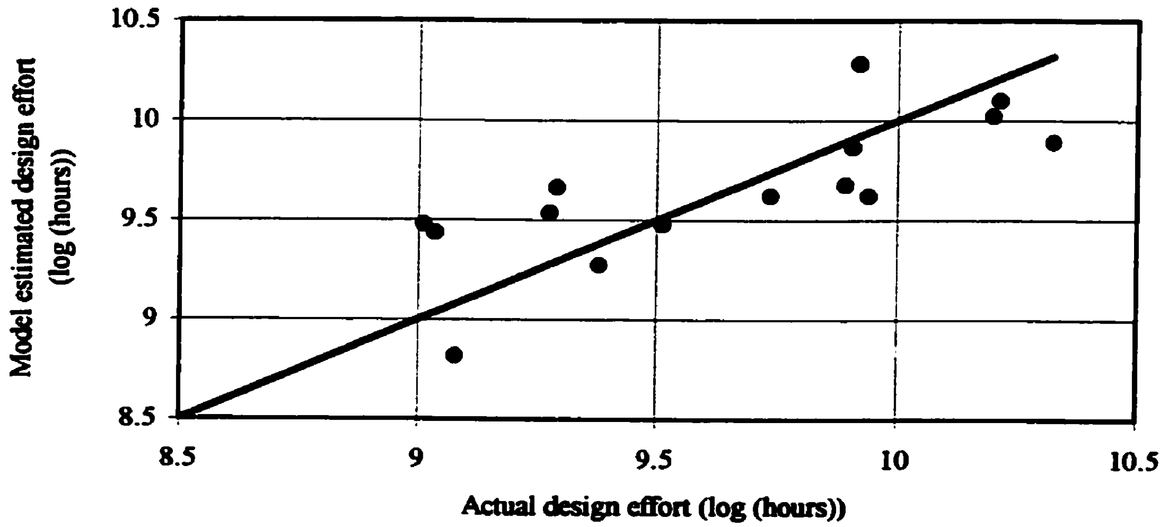



**Figure 5.6** Actual design effort versus single variable model (equation (5.9)) estimates for GE

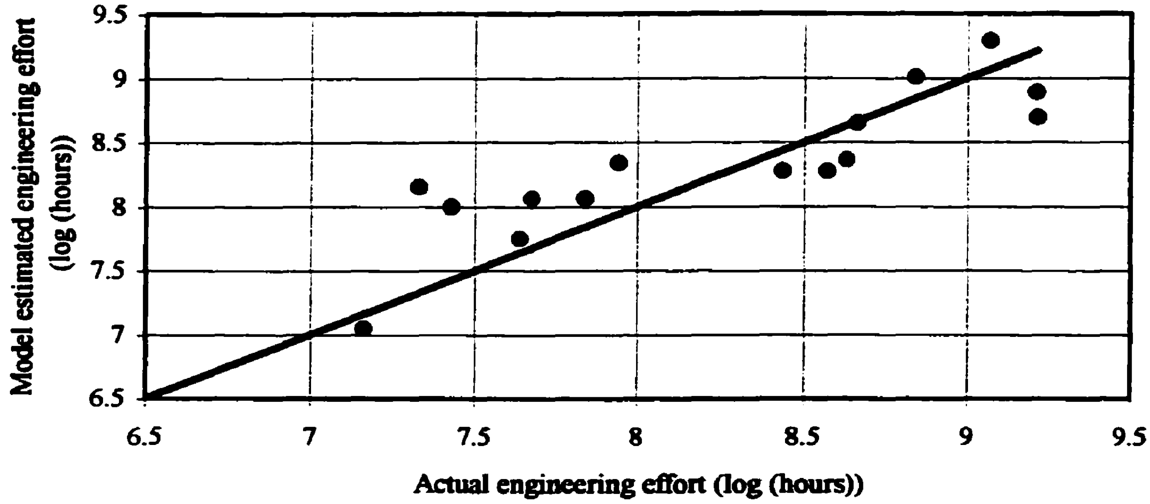



**Figure 5.7** Actual engineering effort versus single variable model (equation (5.10)) estimates for GE

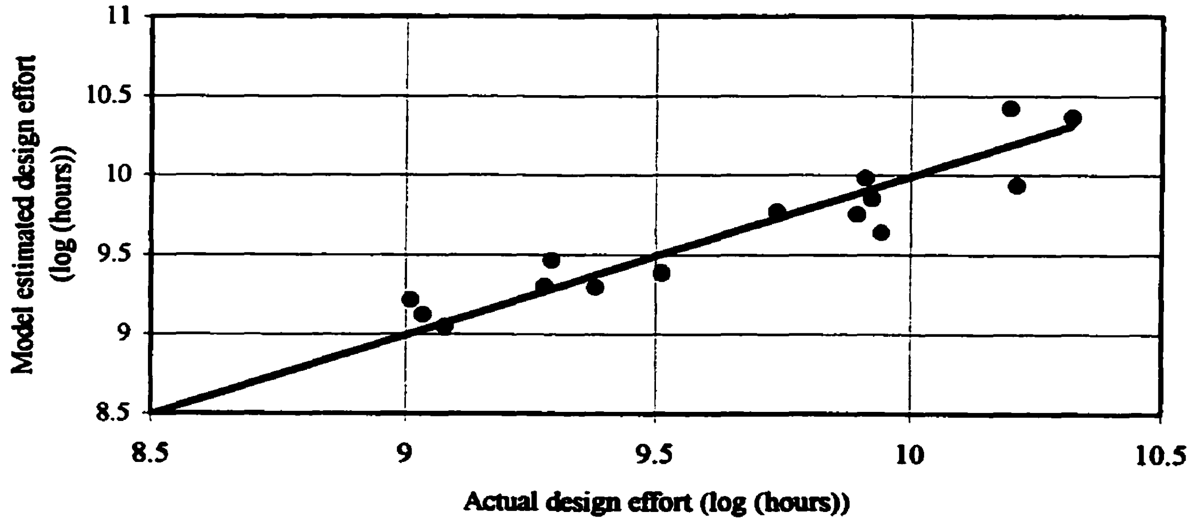



**Figure 5.8** Actual design effort versus multivariable variable model (equation (5.11)) estimates for GE

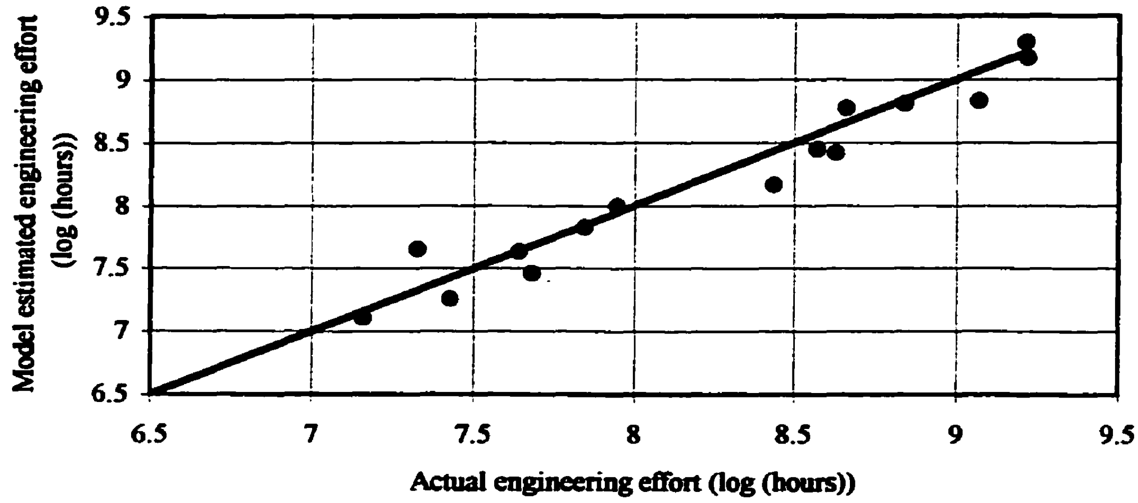



**Figure 5.9** Actual engineering effort versus multivariable variable model (equation (5.12)) estimates for GE

### 5.4.1 Evaluation of the Model for NN

As shown in Table 5.1, the single variable model for NN ($\hat{E} = 24.33PC^{1.39}$ (5.4)) works well as indicated by *MMRE*, *PRED*(.25), and $R^2$ tests. The computed *MMRE* is 12%, and 4 out of 5 cases have error rates less than or equal to 25%, so that *PRED*(.25) = 80%. The computed $R^2$ is 94%.

**Table 5.1** Evaluation of the single variable model (equation (5.4)) for the data from NN

| System number | Product complexity | Actual design effort (hours) | Estimated design effort (hours) | Error (%) |
|---|---|---|---|---|
| 1 | 43 | 4616 | 4536 | 2 |
| 2 | 73 | 8800 | 9466 | 3 |
| 3 | 76 | 7500 | 10011 | 33 |
| 4 | 90 | 11468 | 12663 | 10 |
| 5 | 135 | 25033 | 22249 | 11 |
| | | | | *MMRE* = 12 |
| | | | | $R^2$ (%) = 94 |

### 5.4.2 Evaluation of the Models for CMC

The single variable model for CMC ($\hat{E} = 60.44PC^{1.41}$ (5.5)) does not work well for *MMRE* and *PRED*(.25) tests. As shown in Table 5.2, the computed *MMRE* is 30%, and only 5 out of 12 cases have error rates less than or equal to 25%. In other words, only 42% of the model estimates are within 25% of the actual values. This is most likely due to the existence of other factors influencing design effort. Adding the severity of requirements factor improves the estimation accuracy. As shown in Table 5.3, the computed *MMRE* for the multivariable model ($\hat{E} = 70.65\ PC^{1.15}\ SR^{0.87}$ (5.6)) is 15%, and 92% of the model estimates are within 25% of the actual values. Thus, the improvements in *MMRE* and *PRED*(.25) accomplished by the multivariable model over the single variable model are 50% and 119%, respectively. In spite of the variation using *MMRE* and *PRED*(.25) tests, the two models work well for $R^2$ test, where $R^2$ are 84% and 94%.

**Table 5.2** Evaluation of the single variable model (equation (5.5)) for the data from CMC

| Project number | Product complexity | Actual design effort (hours) | Estimated design effort (hours) | Error (%) |
|---|---|---|---|---|
| 1 | 5 | 951 | 585 | 39 |
| 2 | 7 | 632 | 940 | 49 |
| 3 | 7 | 1103 | 940 | 15 |
| 4 | 11 | 1099 | 1777 | 62 |
| 5 | 11 | 1367 | 1777 | 30 |
| 6 | 11 | 1874 | 1777 | 5 |
| 7 | 11 | 1985 | 1777 | 10 |
| 8 | 15 | 1777 | 2752 | 55 |
| 9 | 19 | 4950 | 3840 | 22 |
| 10 | 22 | 3701 | 4722 | 28 |
| 11 | 24 | 8883 | 5339 | 40 |
| 12 | 34 | 9828 | 8724 | 11 |
| | | | | *MMRE* = 30 |
| | | | | $R^2$ (%) = 84 |

**Table 5.3** Evaluation of the multivariable model (equation (5.6)) for the data from CMC

| Project number | Product complexity | Severity of requirements | Actual design effort (hours) | Estimated design effort (hours) | Error (%) |
|---|---|---|---|---|---|
| 1 | 5 | 2 | 951 | 822 | 14 |
| 2 | 7 | 1 | 632 | 662 | 5 |
| 3 | 7 | 2 | 1103 | 1210 | 10 |
| 4 | 11 | 1 | 1099 | 1114 | 1 |
| 5 | 11 | 2 | 1367 | 2035 | 49 |
| 6 | 11 | 2 | 1874 | 2035 | 9 |
| 7 | 11 | 2 | 1985 | 2035 | 3 |
| 8 | 15 | 1 | 1777 | 1591 | 10 |
| 9 | 19 | 2 | 4950 | 3816 | 23 |
| 10 | 22 | 2 | 3701 | 4516 | 22 |
| 11 | 24 | 3 | 8883 | 7103 | 20 |
| 12 | 34 | 3 | 9828 | 10602 | 8 |
| | | | | | $MMRE = 15$ |
| | | | | | $R^2$ (%) = 94 |

### 5.4.3 Evaluation of a General Model

As indicated previously, an attempt was made to develop a general model capable of estimating design effort for any environment. This concept was tested by combining data from the projects of NN and CMC. As shown in Table 5.4, the single variable model ($\hat{E} = 158.04\ PC^{0.99}$ (5.7)) does not work well for *MMRE* and *PRED*(.25) tests. The computed *MMRE* is 32%, and only 8 out of 17 cases have error rates less than or equal to 25%, so *PRED*(.25) = 47%. As Table 5.5 shows, the multivariable model ($\hat{E} = 151.38\ PC^{0.82}\ SR^{0.84}$ (5.8)) is better than the single variable model with *MMRE* being 23%. However, only 59% of the model estimates are within 25% of the actual values. The two models work well for $R^2$ test. The computed $R^2$ are 83% and 88% for equations (5.7) and (5.8), respectively.

The development of a general model using data from different companies was not possible. This is probably due to the fact that models of two variables do not account for many differences in design environments in the different companies.

**Table 5.4** Evaluation of the single variable model (equation (5.7)) for the combined data from NN and CMC

| Company | Project number | Product complexity | Actual design effort (hours) | Estimated design effort (hours) | Error (%) |
|---|---|---|---|---|---|
| 1 | 1 | 43 | 4616 | 6545 | 42 |
| | 2 | 73 | 8800 | 11052 | 26 |
| | 3 | 76 | 7500 | 11502 | 53 |
| | 4 | 90 | 11468 | 13598 | 19 |
| | 5 | 135 | 25033 | 20314 | 19 |
| 2 | 6 | 5 | 951 | 778 | 18 |
| | 7 | 7 | 632 | 1085 | 72 |
| | 8 | 7 | 1103 | 1085 | 2 |
| | 9 | 11 | 1099 | 1697 | 54 |
| | 10 | 11 | 1367 | 1697 | 24 |
| | 11 | 11 | 1874 | 1697 | 9 |
| | 12 | 11 | 1985 | 1697 | 15 |
| | 13 | 15 | 1777 | 2307 | 30 |
| | 14 | 19 | 4950 | 2916 | 41 |
| | 15 | 22 | 3701 | 3371 | 9 |
| | 16 | 24 | 8883 | 3674 | 59 |
| | 17 | 34 | 9828 | 5187 | 47 |
| | | | | | $MMRE = 32$ |
| | | | | | $R^2$ (%) = 83 |

**Table 5.5** Evaluation of the multivariable variable model (equation (5.8)) for the combined data from NN and CMC

| Company | Project number | Product complexity | Severity of requirements | Actual design effort (hours) | Estimated design effort (hours) | Error (%) |
|---|---|---|---|---|---|---|
| 1 | 1 | 43 | 3 | 4616 | 8323 | 80 |
| | 2 | 73 | 2 | 8800 | 9138 | 4 |
| | 3 | 76 | 2 | 7500 | 9445 | 26 |
| | 4 | 90 | 3 | 11468 | 15252 | 33 |
| | 5 | 135 | 3 | 25033 | 21269 | 15 |
| 2 | 6 | 5 | 2 | 951 | 1014 | 7 |
| | 7 | 7 | 1 | 632 | 747 | 18 |
| | 8 | 7 | 2 | 1103 | 1336 | 21 |
| | 9 | 11 | 1 | 1099 | 1081 | 2 |
| | 10 | 11 | 2 | 1367 | 1936 | 42 |
| | 11 | 11 | 2 | 1874 | 1936 | 3 |
| | 12 | 11 | 2 | 1985 | 1936 | 2 |
| | 13 | 15 | 1 | 1777 | 1394 | 22 |
| | 14 | 19 | 2 | 4950 | 3030 | 39 |
| | 15 | 22 | 2 | 3701 | 3418 | 8 |
| | 16 | 24 | 3 | 8883 | 5160 | 42 |
| | 17 | 34 | 3 | 9828 | 6865 | 30 |
| | | | | | | $MMRE = 23$ |
| | | | | | | $R^2$ (%) = 88 |

### 5.4.4 Evaluation of the Models for GE

The single variable models ($\hat{E} = 3.2\times10^{-8}\, PC^{4.63}$ (5.9) and $\hat{E}_e = 5.9\times10^{-15}\, PC^{7.07}$ (5.10)) do not work well for *MMRE*, *PRED*(.25), and $R^2$ tests. As shown in Tables 5.6 and 5.7, for equations (5.9) and (5.10), the computed *MMRE* are 26% and 35%, the computed *PRED*(.25) are 53% and 60%, and the computed $R^2$ are 54% and 67%, respectively. Adding difficulty to expertise ratio, type of drawings submitted to the customer, and involvement of design partners factors improves the performance of the models. As shown in Tables 5.8 and 5.9, both of the multivariable models ($\hat{E} = 0.12\, PC^2\, DE^{0.41}\, TD^{0.35}\, DP^{0.69}$ (5.11) and $\hat{E}_e = 4.8\times10^{-8}\, PC^{4.29}\, DE^{0.9}\, TD^{0.45}\, DP^{0.68}$ (5.12)) have the same performance with *MMRE* and *PRED*(.25) being 13% and 93%, respectively. Thus, the improvements in *MMRE* and *PRED*(.25) accomplished by equation (5.11) over equation (5.9) are 50% and 75%, respectively. The improvements in *MMRE* and *PRED*(.25) accomplished by equation (5.12) over equation (5.10) are 63% and 55%, respectively.

The computed $R^2$ for equations (5.11) and (5.12) are 81% and 94%, respectively. Thus, the improvements in $R^2$ accomplished by equations (5.11) and (5.12) over corresponding equations (5.9) and (5.10) are 50% and 40%, respectively.

**Table 5.6** Evaluation of the single variable model (equation (5.9)) for the data from GE

| Project number | Product complexity | Actual design effort (hours) | Estimated design effort (hours) | Error (%) |
|---|---|---|---|---|
| 1 | 383 | 20392 | 29198 | 43 |
| 2 | 322 | 8192 | 13077 | 60 |
| 3 | 322 | 13544 | 13077 | 3 |
| 4 | 308 | 11880 | 10645 | 10 |
| 5 | 319 | 8384 | 12523 | 49 |
| 6 | 368 | 27200 | 24268 | 11 |
| 7 | 332 | 20800 | 15067 | 28 |
| 8 | 352 | 30400 | 19753 | 35 |
| 9 | 336 | 19824 | 15926 | 20 |
| 10 | 332 | 16944 | 15067 | 11 |
| 11 | 350 | 20112 | 19239 | 4 |
| 12 | 362 | 26816 | 22489 | 16 |
| 13 | 326 | 10704 | 13847 | 29 |
| 14 | 335 | 10856 | 15708 | 45 |
| 15 | 279 | 8760 | 6734 | 23 |
| | | | | *MMRE* = 26 |
| | | | | $R^2$ (%) = 54 |

### 5.4.5 Parametric Models versus Original Company Estimations

In addition to the above investigations, *MMRE* and *PRED*(.25) of the original estimations made by the three companies shown in Appendix I were compared with those of the developed parametric models which work well for *MMRE*, *PRED*(.25), and $R^2$ tests. The comparison results are summarized in Table 5.10. It can be see from this table that the improvements in estimation accuracy are significant. The improvements in *MMRE* and *PRED*(.25) accomplished by the models over the original estimations ranged from 52-64% and 33-133%, respectively.

**Table 5.7** Evaluation of the single variable model (equation (5.10)) for the data from GE

| Project number | Product complexity | Actual engineering effort (hours) | Estimated engineering effort (hours) | Error (%) |
|---|---|---|---|---|
| 1 | 383 | 8672 | 10816 | 25 |
| 2 | 322 | 2160 | 3172 | 47 |
| 3 | 322 | 2544 | 3172 | 25 |
| 4 | 308 | 2080 | 2317 | 11 |
| 5 | 319 | 1680 | 2969 | 77 |
| 6 | 368 | 6904 | 8154 | 18 |
| 7 | 332 | 4616 | 3938 | 15 |
| 8 | 352 | 10048 | 5955 | 41 |
| 9 | 336 | 5592 | 4286 | 23 |
| 10 | 332 | 5272 | 3938 | 25 |
| 11 | 350 | 5776 | 5720 | 1 |
| 12 | 362 | 10008 | 7260 | 27 |
| 13 | 326 | 1520 | 3462 | 128 |
| 14 | 335 | 2816 | 4197 | 49 |
| 15 | 279 | 1288 | 1152 | 11 |

$MMRE = 35$
$R^2$ (%) = 67

**Table 5.8** Evaluation of the multivariable variable model (equation (5.11)) for data from GE

| Project number | Product complexity | Difficulty to expertise ratio | Type of drawings | Involvement of design partners | Actual design effort (hours) | Estimated design effort (hours) | Error (%) |
|---|---|---|---|---|---|---|---|
| 1 | 383 | 1.2 | 1 | 1 | 20392 | 18969 | 7 |
| 2 | 322 | 0.6 | 1 | 1 | 8192 | 10091 | 23 |
| 3 | 322 | 0.9 | 1 | 1 | 13544 | 11916 | 12 |
| 4 | 308 | 0.9 | 1 | 1 | 11880 | 10902 | 8 |
| 5 | 319 | 0.5 | 1 | 1 | 8384 | 9191 | 10 |
| 6 | 368 | 1.0 | 2 | 1 | 27200 | 20713 | 24 |
| 7 | 332 | 0.8 | 2 | 1 | 20800 | 15385 | 26 |
| 8 | 352 | 1.1 | 2 | 2 | 30400 | 31791 | 5 |
| 9 | 336 | 1.0 | 2 | 1 | 19824 | 17267 | 13 |
| 10 | 332 | 1.1 | 2 | 1 | 16944 | 17530 | 3 |
| 11 | 350 | 1.0 | 3 | 1 | 20112 | 21593 | 7 |
| 12 | 362 | 1.1 | 2 | 2 | 26816 | 33623 | 25 |
| 13 | 326 | 0.7 | 1 | 1 | 10704 | 11018 | 3 |
| 14 | 335 | 0.9 | 1 | 1 | 10856 | 12898 | 19 |
| 15 | 279 | 0.8 | 1 | 1 | 8760 | 8524 | 3 |

$MMRE = 13$
$R^2$ (%) = 81

**Table 5.9** Evaluation of the multivariable variable model (equation (5.12)) for data from GE

| Project number | Product complexity | Difficulty to expertise ratio | Type of drawings | Involvement of design partners | Actual engineering effort (hours) | Estimated design effort (hours) | Error (%) |
|---|---|---|---|---|---|---|---|
| 1 | 383 | 1.2 | 1 | 1 | 8672 | 6830 | 21 |
| 2 | 322 | 0.6 | 1 | 1 | 2160 | 1739 | 19 |
| 3 | 322 | 0.9 | 1 | 1 | 2544 | 2505 | 2 |
| 4 | 308 | 0.9 | 1 | 1 | 2080 | 2070 | 0 |
| 5 | 319 | 0.5 | 1 | 1 | 1680 | 1418 | 16 |
| 6 | 368 | 1.0 | 2 | 1 | 6904 | 6671 | 3 |
| 7 | 332 | 0.8 | 2 | 1 | 4616 | 3509 | 24 |
| 8 | 352 | 1.1 | 2 | 2 | 10048 | 9623 | 4 |
| 9 | 336 | 1.0 | 2 | 1 | 5592 | 4515 | 19 |
| 10 | 332 | 1.1 | 2 | 1 | 5272 | 4673 | 11 |
| 11 | 350 | 1.0 | 3 | 1 | 5776 | 6457 | 12 |
| 12 | 362 | 1.1 | 2 | 2 | 10008 | 10852 | 8 |
| 13 | 326 | 0.7 | 1 | 1 | 1520 | 2106 | 39 |
| 14 | 335 | 0.9 | 1 | 1 | 2816 | 2968 | 5 |
| 15 | 279 | 0.8 | 1 | 1 | 1288 | 1218 | 5 |

*MMRE* = 13
$R^2$ (%) = 94

**Table 5.10** Parametric models versus original company estimations

| Company | Model | Model estimations | | Company original estimations | | Improvement | |
|---|---|---|---|---|---|---|---|
| | | *MMRE* (%) | *PRED*(.25) (%) | *MMRE* (%) | *PRED*(.25) (%) | *MMRE* (%) | *PRED*(.25) (%) |
| NN | $\hat{E} = 24.33PC^{1.39}$ (5.4) | 12 | 80 | 33 | 60 | 64 | 33 |
| CMC | $\hat{E} = 70.65PC^{1.15}SR^{0.87}$ (5.6) | 15 | 92 | 41 | 42 | 63 | 119 |
| GE | $\hat{E} = 0.12\, PC^{2}\, DE^{0.41}\, TD^{0.35}\, DP^{0.69}$ (5.11) | 13 | 93 | 27 | 53 | 52 | 75 |
| | $\hat{E}_e = 4.8\times10^{-8}\, PC^{4.29}\, DE^{0.9}\, TD^{0.45}\, DP^{0.68}$ (5.12) | 13 | 93 | 31 | 40 | 58 | 133 |

## 5.4 Estimation Charts

One advantage of parametric estimation models is that they can be represented by simple charts, which can be used to obtain quick estimates for design effort. For demonstration, the charts shown in Figures 5.10-5.17 were prepared using equations $\hat{E} = 24.33PC^{1.39}$ (5.4), $\hat{E} = 70.65PC^{1.15}\ SR^{.87}$ (5.6), and $\hat{E} = 0.12\ PC^{2}\ DE^{0.41}\ TD^{0.35}\ DP^{0.69}$ (5.11). Demonstrative examples on how to use these charts are presented below.

<u>Example 5.1</u>

The chart in Figure 5.10 gives the relationship between product complexity, $PC$, (horizontal axis) and estimated design effort in hours, $\hat{E}$, (vertical axis) for NN projects using equation (5.4). For example, a project will require approximately 12500 hours of design effort, if its corresponding $PC$ is 90.

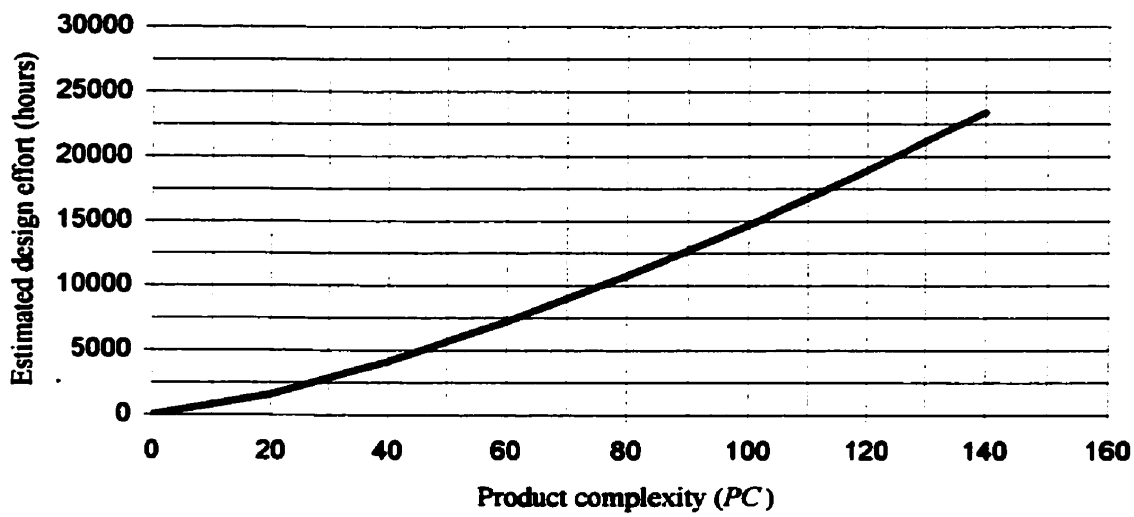



**Figure 5.10** Estimation chart for NN using equation (5.4)

<u>Example 5.2</u>

The chart in Figure 5.11 was obtained for CMC using equation (5.6). Each curve in this chart represents the relationship between product complexity, $PC$, (horizontal axis) and estimated design effort in hours, $\hat{E}$, (vertical axis) at a certain level of severity of requirements, *SR*. For example, a project will require approximately 5000 hours of design effort, if $PC = 24$ and $SR = 2$.

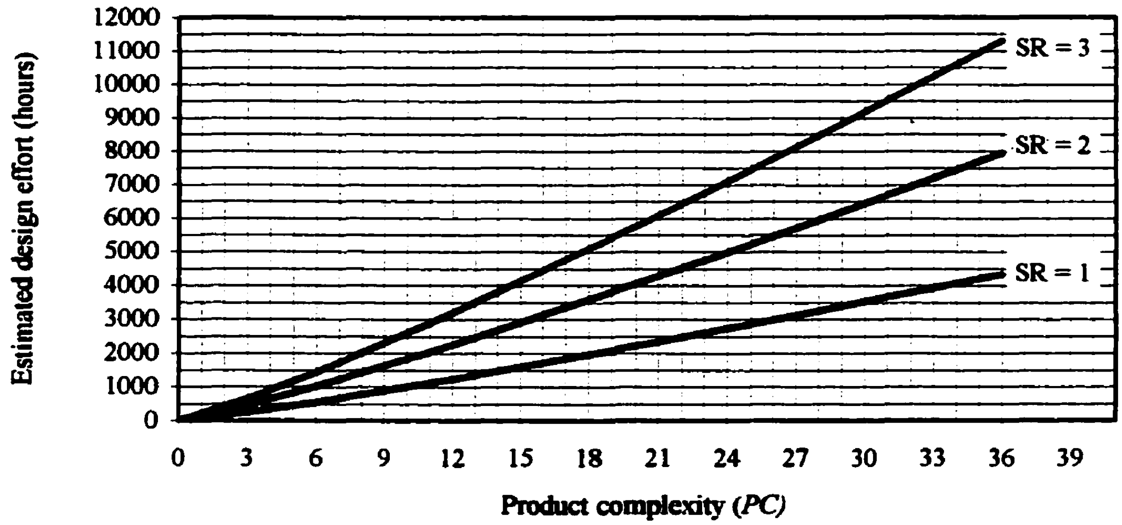



**Figure 5.11** Estimation chart for CMC using equation (5.6) where *PC* and *SR* are varied

<u>Example 5.3</u>

The chart in Figure 5.12 was obtained for GE using equation (5.11). Each curve in this chart represents the relationship between product complexity, *PC*, (horizontal axis) and estimated design effort in hours, $\hat{E}$, (vertical axis) at a certain value of difficulty to expertise ratio, *DE*, where involvement of design partners, $DP = 1$, and type of drawings, $TD = 1$. For example, a project will require approximately 9100 hours of design effort, if its corresponding $PC = 318$, $DE = 0.5$, $TD = 1$, and $DP = 1$.

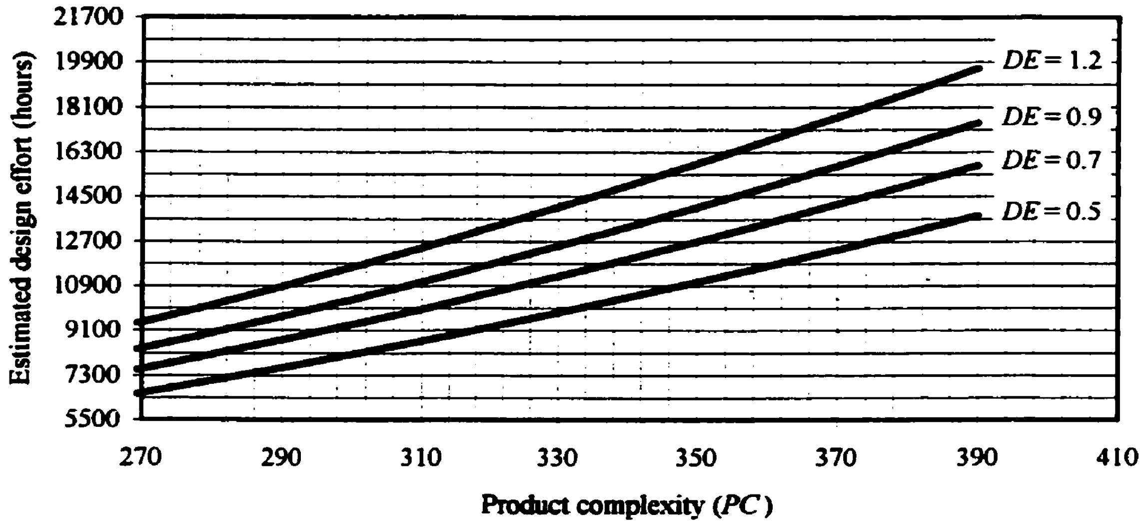



**Figure 5.12** Estimation chart for GE using equation (5.11) where *TD* = 1, *DP* = 1, *DE* and *PC* are varied

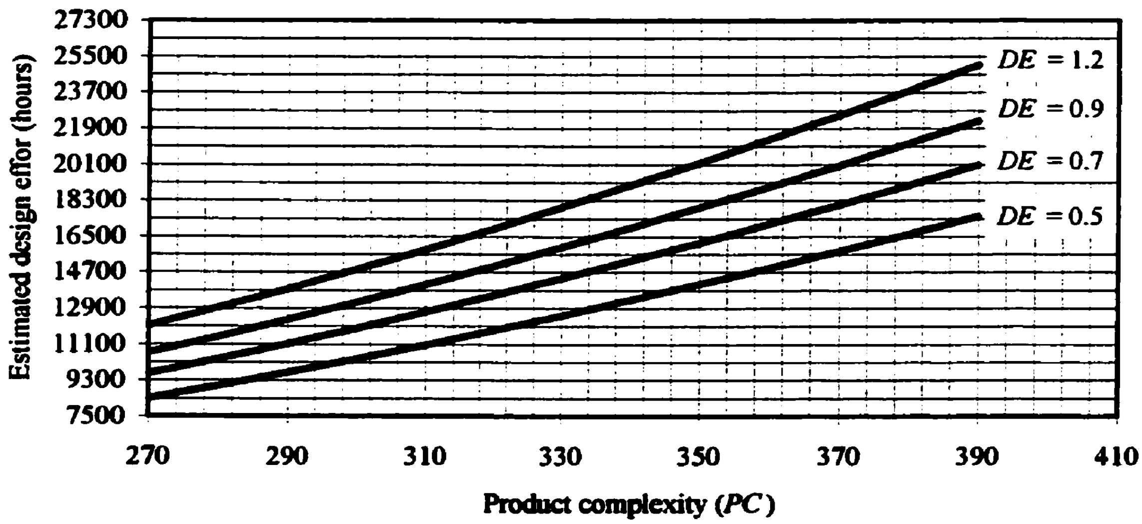



**Figure 5.13** Estimation chart for GE using equation (5.11) where *TD* = 2, *DP* = 1, *DE* and *PC* are varied

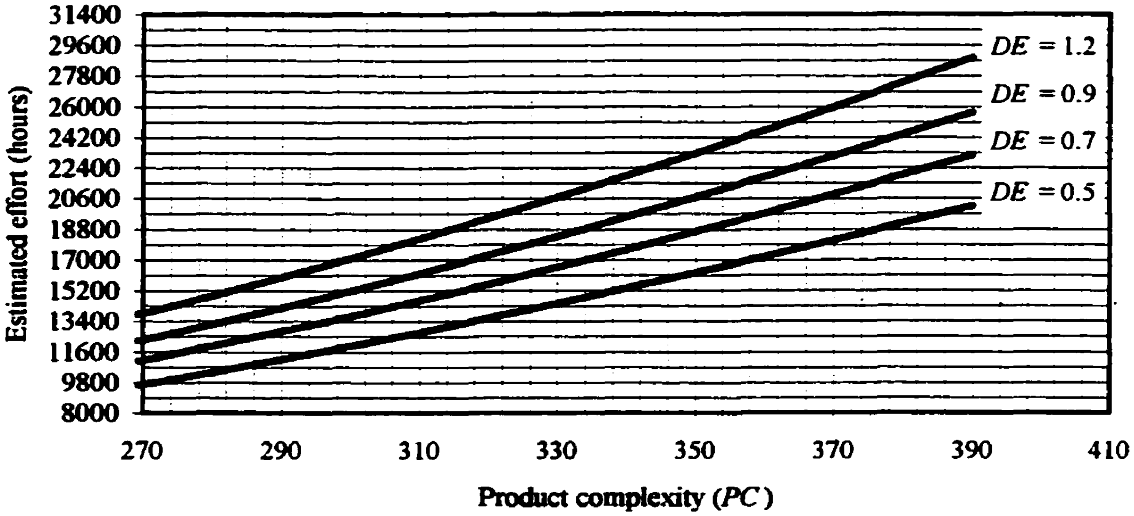



**Figure 5.14** Estimation chart for GE using equation (5.11) where *TD* = 3, *DP* = 1, *DE* and *PC* are varied

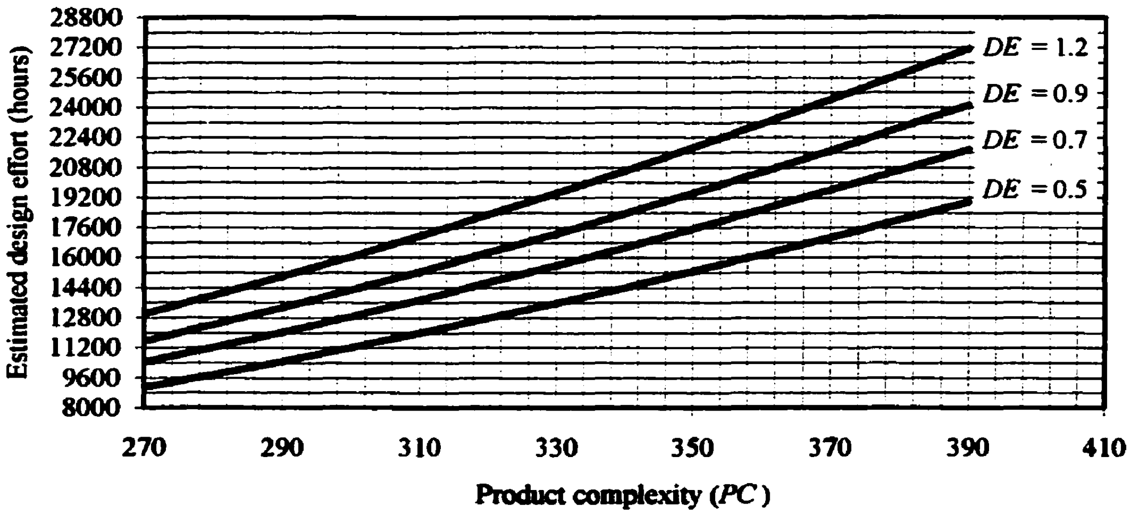



**Figure 5.15** Estimation chart for GE using equation (5.11) where *TD* = 1, *DP* = 2, *DE* and *PC* are varied

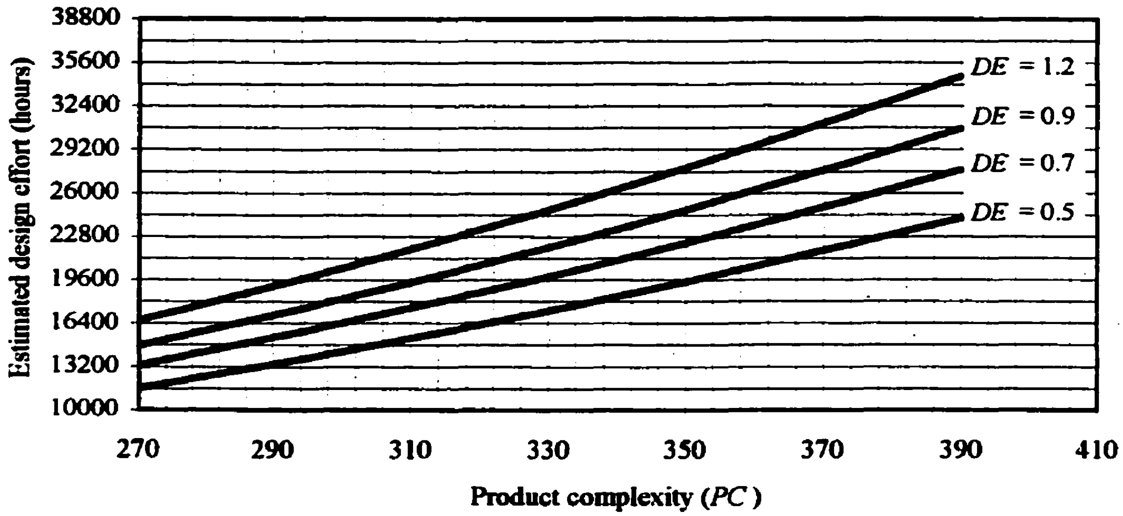



**Figure 5.16** Estimation chart for GE using equation (5.11) where $TD$ = 2, $DP$ = 2, $DE$ and $PC$ are varied

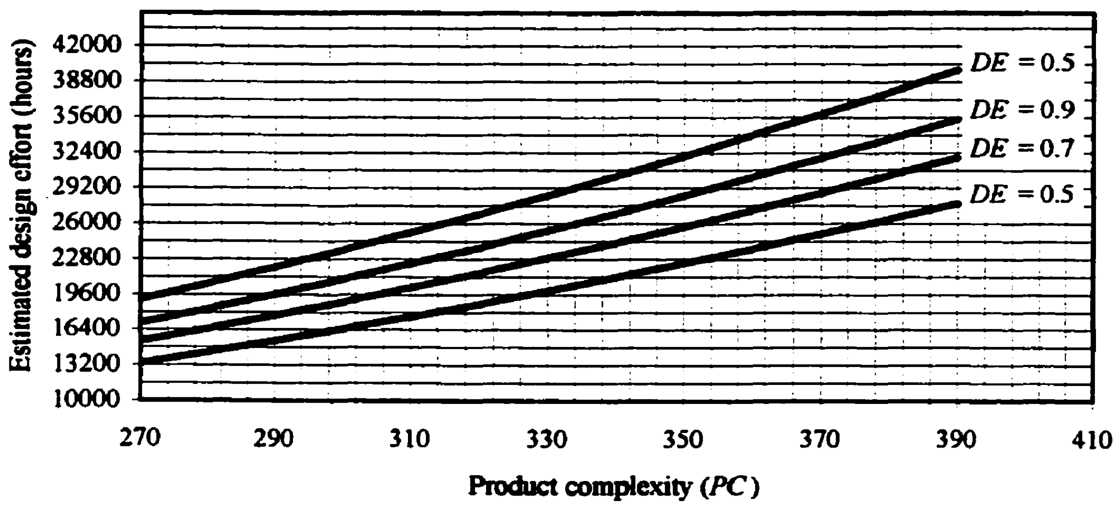



**Figure 5.17** Estimation chart for GE using equation (5.11) where $TD$ = 3, $DP$ = 2, $DE$ and $PC$ are varied

## 5.5 Summary

Using traditional regression analysis, parametric models were developed from data from NN, CMC, and GE. The general form of the models is defined by equation (5.1). Generally, the models performed well according to a number of accuracy tests.

A comparison was made between the estimations obtained using the parametric models and those made by the companies. The results indicated that the parametric models were better predictors than the original design managers using expert judgement.

The use of one general model to estimate design effort for different companies was tried by constructing models using combined data from NN and CMC. The results indicated that its performance was not as good as the models that were developed from data from single environments. This is probably due to the heterogeneity of the products and/or to the existence of some differences in the environments of the two companies which was not taken into account by the models.

The developed models are useful not only for estimating total design effort, but also for estimating effort for a subset of the total effort, such as engineering effort, as was done for GE.

Similar to software estimation models, more than 50% of variation in estimating effort can be explained by product complexity. This confirms that product complexity is the dominant parameter in estimating design effort spent during a project.

# Chapter 6

# Artificial Neural Network Models

One disadvantage of the parametric models described in the previous chapter is that the form of the regression equation needs to be known a priori or guessed. This means that the regression is constrained to yield a best fit for the specified form of the equation. If the specified form is a poor guess, this constraint can be serious (Specht, 1991). Thus, there is a need for a method that can be used when it is difficult to define the form that best fits historical data. For this purpose, this chapter describes the application of artificial neural networks to design effort estimation, where a priori assumption about the equation form is not required (Bashir and Thomson, 1999e).

## 6.1 Artificial Neural Networks (ANNs)

Since the 1950's, extensive research has been carried out in the area of artificial intelligence (AI). Such research has led to the emergence of many techniques that simulate the ways of problem solving by humans (Krishnamoorthy and Rajeev, 1996). ANNs is one of such techniques that attempts to mimic biological neural systems both in functionality and in structure. Functionality includes pattern classifications or predictions based on past experience (Wasserman, 1989). Neural networks process information through the interaction of a large number of simple processing units known as neurons. Each neuron, in its simplest form, receives a number of input signals; then, each input is multiplied by a weight, and all the weighted inputs are summed up to determine the activation level of the processing unit. This activation is converted into an output signal by a transfer function (Wasserman, 1989). A neural network is normally constructed by arranging processing units in a number of layers. As shown in Figure 6.1, a simple neural network consists of input layer, hidden layer, and output layer. If a neuron receives data from outside of the network, it is considered to be in the input layer. If it contains the predictions or classifications, the neuron is considered in the output layer. Any neuron in between the input and output layers is considered in the hidden layer.

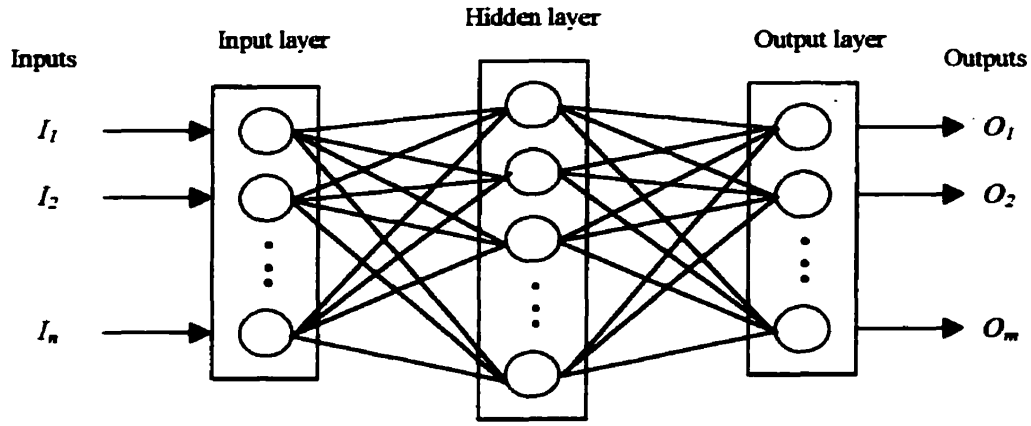



**Figure 6.1** Simple artificial neural network architecture

ANNs can be classified into two basic types: supervised and unsupervised. A supervised network makes predictions or classifications after it is fed with a number of correct classifications or predictions from which it can learn. An unsupervised network makes classifications without being shown in advance how to categorize.

ANNs have many unique characteristics.

- They do not require a priori assumptions about the equation's form.
- Models with multiple outputs can be built using ANNs.
- ANNs are able to function well with noisy or slightly incorrect data.

On the other hand, the process of developing an artificial neural network model is not straightforward. It requires some trial and error to select the proper architecture and to set its parameters.

## 6.2 ANNs for Estimating Design Effort

Developing an ANN model requires a careful selection of the factors that have predictive relationships to design effort. Inclusion of one or more of the above factors depends on the characteristics of the historical projects in the data set. As was indicated in Chapter 3, from the characteristics of the historical projects that were collected from CMC and GE the following factors were identified as predictors for design effort.

<u>CMC</u>

- Product complexity
- Severity of requirements

<u>GE</u>

- Product complexity
- Technical difficulty
- Team expertise
- Type of drawings submitted to the customer
- Involvement of design partners

As indicated in Chapter 4, product complexity was estimated by the metric, *PC*, defined by equation (4.1) for the data collected from CMC, and by the metric defined by equation (4.2) for the data collected from GE. The other factors were assigned numerical values as described in Chapter 3.

Having identified the factors, a paradigm must be selected. Several neural networks with distinct capabilities have been developed. These include Perceptron (Rosenblatt, 1961), back propagation (Rumelhart et al., 1986), counter propagation (Hecht-Nielsen, 1987, 1988), Boltzmann machine (Hinton and Sejnowski, 1986), Hopfield (Hopfield, 1982), BAM (Kosko, 1987), ART (Carpenter and Grossberg, 1987), probabilistic neural network (Specht, 1988), and general regression neural network (Specht, 1991). More details about the advantages and disadvantages of these paradigms are not presented. They may be referred to in the relevant literature. Special attention is directed to Bailey and Thompson

(1990). In this thesis, because of the small sample size, the General Regression Neural Network (GRNN) was chosen in preference to other paradigms.

### 6.2.1 General Regression Neural Network (GRNN)

GRNN is a type of supervised network. It has the ability to train quickly on sparse data without getting trapped by locally optimal solutions (Specht, 1991).

#### 6.2.1.1 Mathematical Background

It is known that the conditional mean of a random variable, $y_o$, can be defined by equation (6.1).

$$E[y_o / X] = \frac{\int_{-\infty}^{\infty} y_o\, f(X, y_o)\, dy_o}{\int_{-\infty}^{\infty} f(X, y_o)\, dy_o} \tag{6.1}$$

where:

$E[y_o / X]$ = *the conditional mean of the dependant variable* $y_o$ *given the independent vector* $X$

$f(X, y_o)$ = *joint probability density function (pdf)*

Equation (6.1) indicates that if the joint pdf is known, then, the conditional mean of $y_o$ given $X$ can be computed. In practice, however, the pdf is usually unknown and can only be estimated from a sample of observations. Parzen (1962) presented a method for estimating a univariate pdf, which was extended later to a multivariate case by Cacoullos (1966). Using this method, Specht (1991) showed that the expected value, $\hat{Y}(X)$, can be computed using equation (6.2).

$$\hat{Y}(X) = \frac{\sum_{i=1}^{n} Y_i\, e^{-\frac{D_i^2}{2\sigma^2}}}{\sum_{i=1}^{n} e^{-\frac{D_i^2}{2\sigma^2}}} \tag{6.2}$$

where:

$D_i^2 = (X - X_i)^T (X - X_i)$

$\sigma$ = *smoothing factor*

$n$ = *sample size*

Equation (6.2) is the fundamental equation of the GRNN. The resulting regression procedure is implemented via three-layer network architecture with one hidden layer. As shown in Figure 6.2, the hidden layer consists of pattern units and summation units[14]. The number of neurons in the hidden layer is usually equal to the number of patterns in the training set.

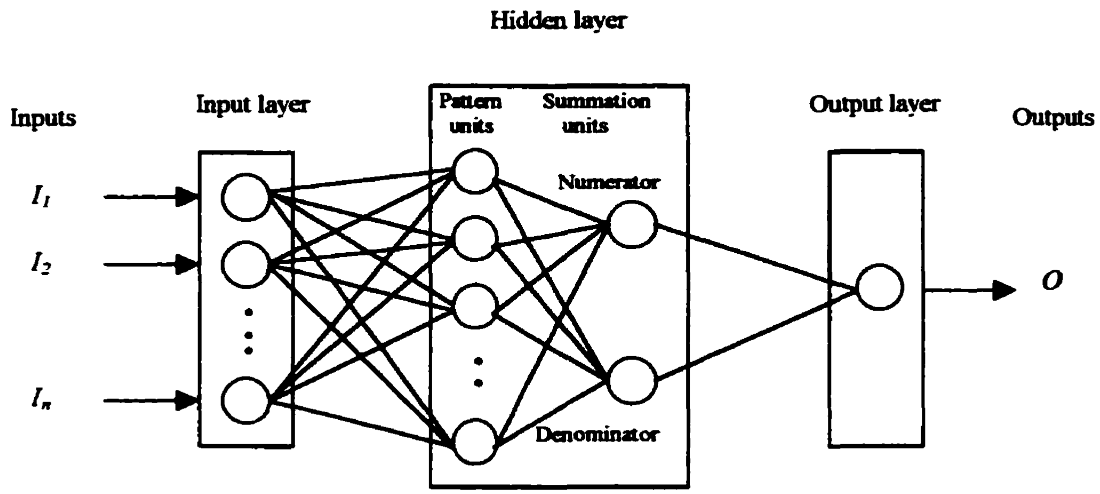



**Figure 6.2** General neural network (GRNN) architecture

### 6.2.2 Development of the GRNN Estimation Models

Using NeuroShell 2 developed by Ward Systems Group, Inc. as a software tool, GRNN models were designed and trained to predict the design effort for the data collected from CMC and GE. The input layer of the model developed for CMC had two neurons: product

[14] For multivariate prediction, a numerator summation unit is needed for each dependent variable.

complexity and severity of requirements, while the input layer of the model developed for GE had four neurons: product complexity, technical difficulty to team expertise ratio, type of drawing submitted to the customer, and involvement of design partners. The output layer for each model had a single neuron, which represents the estimated effort.

The number of neurons in the hidden layer were set to be 9 and 11 for the models developed for CMC and GE models, respectively. Each data set was split randomly into a learning set and a test set. For the data from CMC, the learning set consisted of projects 1-3, 5-9, and 13. For the data from GE, the learning set consisted of projects 1, 2, 5, 6, 8, 9, and 11-15. The remaining projects from each data set were used as a test set to evaluate the performance of the trained networks. The smoothing factors were varied to improve the network performance. Smoothing factors, $\sigma$, of 0.019 and 0.297 for the constructed models for CMC and GE, respectively, gave good results.

## 6.3 Performance Evaluation Results

The GRNN models for CMC and GE work well as indicated by the mean magnitude of relative error (*MMRE*), prediction at 25% (*PRED*(.25)), and the model's coefficient of multiple determination ($R^2$) tests.

As shown in Table 6.1, the computed *MMRE*, *PRED*(.25), and $R^2$ for CMC are 14%, 75% and 99%, respectively. Furthermore, GRNN estimations for the three projects in the test set, are 1099, 3701, and 8883 hours, respectively, resulting in no estimation errors.

Table 6.2 shows that the computed *MMRE*, *PRED*(.25), and $R^2$ for CMC are 13%, 93% and 81%, respectively. The actual effort spent on the four projects in the test set were 13544, 11880, 20800, and 16944 hours. The GRNN model estimated 11977, 11603, 17359, and 19726 for the four projects respectively, resulting in error estimations of 12%, 2%, 17%, and 16%.

**Table 6.1** Evaluation of the GRNN model for the data from CMC

| Project number | Actual design effort (hours) | Estimated design effort (hours) | Error (%) | Category |
|---|---|---|---|---|
| 1 | 951 | 951 | 0 | Training |
| 2 | 632 | 1099 | 74 | Training |
| 3 | 1103 | 951 | 14 | Training |
| 4 | 1099 | 1099 | 0 | test |
| 5 | 1367 | 1742 | 27 | Training |
| 6 | 1874 | 1742 | 7 | Training |
| 7 | 1985 | 1742 | 12 | Training |
| 8 | 1777 | 1099 | 38 | Training |
| 9 | 4950 | 4950 | 0 | Training |
| 10 | 3701 | 3701 | 0 | test |
| 11 | 8883 | 8883 | 0 | test |
| 12 | 9828 | 9828 | 0 | Training |
| | | | $MMRE = 14$ | |
| | | | $R^2$ (%) = 99 | |

**Table 6.2** Evaluation of the GRNN model for the data from GE

| Project number | Actual design effort (hours) | Estimated design effort (hours) | Error (%) | Category |
|---|---|---|---|---|
| 1 | 20392 | 19856 | 3 | Training |
| 2 | 8192 | 10244 | 25 | Training |
| 3 | 13544 | 11977 | 12 | test |
| 4 | 11880 | 11603 | 2 | test |
| 5 | 8384 | 9649 | 15 | Training |
| 6 | 27200 | 22664 | 17 | Training |
| 7 | 20800 | 17359 | 17 | test |
| 8 | 30400 | 21858 | 28 | Training |
| 9 | 19824 | 19431 | 2 | Training |
| 10 | 16944 | 19726 | 16 | test |
| 11 | 20112 | 21210 | 5 | Training |
| 12 | 26816 | 22726 | 15 | Training |
| 13 | 10704 | 10935 | 2 | Training |
| 14 | 10856 | 12414 | 14 | Training |
| 15 | 8760 | 10479 | 20 | Training |
| | | | $MMRE = 13$ | |
| | | | $R^2$ (%) = 81 | |

## 6.4 GRNN Versus Parametric Models

To compare the performance of the parametric models described in the previous chapter and the GRNN models, a parametric model was established for each company using the same training set used by the corresponding GRNN model. The constructed models for CMC and GE are defined by equations (6.3) and (6.4), respectively.

$$\hat{E} = 71.23\, PC^{1.17}\, SR^{0.70} \tag{6.3}$$

$$\hat{E} = 5.14\, PC^{1.36}\, DE^{0.77}\, TD^{0.23}\, DP^{0.58} \tag{6.4}$$

As shown in Table 6.3, the parametric model for CMC (equation (6.3)) has a mean magnitude of relative error (*MMRE*) of 14%, and 9 out of 12 of the model estimates are within 25% of the actual values, so *PRED*(.25) = 75%. The computed coefficient of multiple determination ($R^2$) is 92%. The estimated efforts for three projects in the test set are 1269, 1150, and 4725 hours, resulting in error estimations of 7%, 16%, and 29%.

**Table 6.3** Evaluation of the parametric model for the data from CMC

| Project number | Actual design effort (hours) | Estimated design effort (hours) | Error (%) | Category |
|---|---|---|---|---|
| 1 | 951 | 761 | 20 | Training |
| 2 | 632 | 694 | 10 | Training |
| 3 | 1103 | 1128 | 2 | Training |
| 4 | 1099 | 1178 | 7 | test |
| 5 | 1367 | 1913 | 40 | Training |
| 6 | 1874 | 1913 | 2 | Training |
| 7 | 1985 | 1913 | 4 | Training |
| 8 | 1777 | 1693 | 5 | Training |
| 9 | 4950 | 3627 | 27 | Training |
| 10 | 3701 | 4305 | 16 | test |
| 11 | 8883 | 6331 | 29 | test |
| 12 | 9828 | 9517 | 3 | Training |
| | | | *MMRE* = 14 | |
| | | | $R^2$ (%) = 92 | |

The parametric model for GE (equation (6.4) works well as indicated by the mean magnitude of relative error (*MMRE*), prediction at 25% (*PRED*(.25)), and the model's coefficient of multiple determination ($R^2$) tests. As shown in Table 6.4, the computed *MMRE*, *PRED*(.25), and $R^2$ are 11%, 87%, and 79%, respectively. Moreover, the estimation errors for the four projects in the test set are 10%, 3%, and 34%, and 3%.

The above comparisons between the GRNN and the parametric models are summarized in Table 6.5. The general conclusion that can be drawn from this table is that the performance of the GRNN and the parametric models are comparable in terms of *MMRE*, *PRED*(.25), and $R^2$.

**Table 6.4** Evaluation of the parametric model for the data from GE

| Project number | Actual design effort (hours) | Estimated design effort (hours) | Error (%) | Category |
|---|---|---|---|---|
| 1 | 20392 | 19279 | 5 | Training |
| 2 | 8192 | 8929 | 9 | Training |
| 3 | 13544 | 12202 | 10 | test |
| 4 | 11880 | 11486 | 3 | test |
| 5 | 8384 | 7662 | 9 | Training |
| 6 | 27200 | 18610 | 32 | Training |
| 7 | 20800 | 13625 | 34 | test |
| 8 | 30400 | 28182 | 7 | Training |
| 9 | 19824 | 16445 | 17 | Training |
| 10 | 16944 | 17411 | 3 | test |
| 11 | 20112 | 19082 | 5 | Training |
| 12 | 26816 | 29276 | 9 | Training |
| 13 | 10704 | 10225 | 4 | Training |
| 14 | 10856 | 12876 | 19 | Training |
| 15 | 8760 | 9170 | 5 | Training |
| | | | *MMRE* = 11<br>$R^2$ (%) = 79 | |

**Table 6.5** GRNN versus parametric models

| Criterion | CMC | | GE | |
|---|---|---|---|---|
| | GRNN | Parametric | GRNN | Parametric |
| The mean magnitude of relative error, *MMRE* (%) | 14 | 14 | 13 | 11 |
| Prediction at 25%, *PRED*(.25) | 75 | 75 | 93 | 87 |
| The coefficient of multiple determination, $R^2$ | 99 | 92 | 81 | 79 |
| Estimation errors (%) for the projects in test sets | 0, 0, 0 | 7, 16, 29 | 12, 2, 17, 16 | 10, 3, 34, 3 |

## 6.5 Summary

Unlike traditional regression analysis, a general regression neural network (GRNN) is based on the probability density function of the observed data rather than on a presumed function. In this chapter, general regression neural network (GRNN) estimation models were constructed for the data from CMC and GE. The input variables were the same as those used by the multivariable parametric estimation models described in the previous chapter.

The results clearly show that the artificial neural networks can be considered as a good tool for estimating design effort. Within the limited data sets, the developed models produced good results. A comparison between the GRNN estimation models with parametric models based on traditional regression analysis showed that both models had about the same accuracy. This conclusion should not be generalized; some models may perform well on certain data, others may not. However, the sole purpose of this work was to demonstrate the capabilities of ANNs as an alternative method for estimating design effort. ANNs can be more practically utilized for cases where a mathematical relationship is not easily established.

# Chapter 7

# An Analogy-Based Model

The parametric and the artificial neural network estimation models described in the previous two chapters cannot be constructed, unless there are patterns of relationships between design effort as the dependent variable and one or more factors as independent variables. In addition, they do not take into account unusual situations; therefore, they are only useful in a relatively constant environment. Thus, there is a need for a method that gives the estimator a means of prediction in unusual situations and/or when parametric or artificial neural network models are difficult to obtain for estimation. For this purpose, this chapter proposes an analogy-based model for estimating design effort. The implementation of the proposed model depends on knowledge of the productivity of reference projects, the size of the upcoming project, and the understanding of the factors that affected the productivity of the reference projects as well as those that will affect the upcoming project's productivity.

## 7.1 Development of the Model

Decision-makers informally use previous cases to make decisions (Ross, 1986). This technique is also applicable to effort estimation (Hughes, 1996). For example, if a new project is believed to be 15% more complex than a previous project, then, the estimate of effort is increased by 15%. In an attempt to formalize this phenomenon, an analogy-based model for estimating design effort is described below.

Since productivity is defined as the ratio of outputs generated from a system to the inputs provided to create the outputs, then, it is possible to measure the productivity of a reference project and use it as a predictor of future productivity, and thus, the expected effort of an upcoming project. The model computations involve the following steps (Bashir and Thomson, 1999d):

1) selection of reference projects,
2) computation of the productivity of the reference projects,
3) identification of the major factors that affected the productivity of the reference projects as well as those that will affect the upcoming project's productivity,
4) computation of the multipliers that capture the effect of each factor on productivity,
5) estimation of the upcoming project's productivity, and
6) estimation of the design effort for the upcoming project.

The above six steps are described below.

### 7.1.1 Selection of Reference Projects

Select a set of completed projects. The selection should be restricted to those that have a high degree of similarity with the upcoming project in terms of factors that influence them. Keeping the number of different factors low helps to minimize the number of required estimates, and therefore, to maximize the overall accuracy of the estimation.

### 7.1.2 Computation of the Productivity of the Reference Projects

Compute the productivity of each reference project, $P_r$, using equation (7.1).

$$P_r = \frac{O_r}{E_r} \tag{7.1}$$

where:
$O_r$ = *the output of reference project r*
$E_r$ = *the input of reference project r*

The input, $E_r$, is defined as the number of man-months which has been spent by designers including project managers on design activities in the time covering the period between the end of the feasibility study and the release of final detailed drawings to manufacturing. While measuring the input is straightforward, measurement of the output, $O_r$, is an elusive concept and difficult to gauge. This is because of the often non-homogeneous and intangible nature of the output. However, because the inherent objective of a design project is to design a product that delivers certain functions, it is proposed to measure the output, $O_r$, by the product complexity metric, *PC*, defined by equations (4.1) and (4.2).

### 7.1.3 Identification of the Major Factors

Identify the factors that explain why some projects have higher or lesser productivity than others. As indicated in Chapter 3, these factors should be identified among more than a hundred factors, which influence different aspects of the design process. However, it should be noted that it is not necessary to consider factors that are constant for the upcoming project and the reference projects. For example, if no computer tools were used in any of the projects, then, there is no need to consider the automated design tools factor. Thus, the number of factors to be considered depends on the similarity among the upcoming project and the reference projects. The higher the similarity, the less number of factors need to be included.

### 7.1.4 Computation of the Multipliers

After a set of factors is identified, it is necessary to compute the multipliers that capture the effect of each factor on productivity. The multipliers can be supplied subjectively by design managers directly. However, to minimize judgmental error, and thus, maximize estimation accuracy, it is proposed to implement an eigenvector approach (Saaty, 1980). In addition to its potential to derive better estimates, the eigenvector approach provides a measure of consistency that is not available in a direct estimation method.

The major steps in determining the multipliers that capture the effect of each factor on productivity are the following:

a) To ensure a certain level of consistency, develop a rating system. One possible rating system is given in Table 7.1. Given two projects, the highest value 9 is assigned when the influence of the factor under consideration on the first project's productivity was/will be extremely severe as compared with the second. The value 1 is assigned when the factor had/will have equal influence on the productivity of the two projects. Reciprocals are assigned to reflect dominance of the second project as compared with the first in terms of the influence of the factor on productivity.

b) Use the rating scale described in Table 7.1, to construct a matrix of pairwise comparisons, $A$. This matrix contains pairwise comparisons between the projects in terms of the relative influences of the factor under consideration on productivity. As shown in Table 7.2, the matrix requires $s \times s$ entries, where $s$ is the number of projects (an upcoming project and a set of reference projects). However, since the comparisons are reciprocal, i.e., $a_{ij} = \frac{1}{a_{ji}}$ for all $i$, $j = 1, 2, \ldots, s$, only $\frac{s(s-1)}{2}$ of the comparisons need to be made.

**Table 7.1** Influence rating scale

| Numerical values | Definition |
|---|---|
| 1 | Equal influence |
| 3 | Slightly more influence |
| 5 | More influence |
| 7 | Severe influence |
| 9 | Extremely severe influence |
| 2, 4, 6, 8 | Intermediate values to reflect compromise |

**Table 7.2** Pairwise comparison matrix

| Project number | 1 | 2 | ... | ... | $s$ |
|---|---|---|---|---|---|
| 1 | $a_{11}$ | $a_{12}$ | ... | ... | $a_{1s}$ |
| 2 | $a_{21}$ | $a_{22}$ | ... | ... | $a_{2s}$ |
| . | . | . | . | . | . |
| . | . | . | . | . | . |
| $s$ | $a_{s1}$ | $a_{s2}$ | ... | ... | $a_{ss}$ |

c) Compute the principal right eigenvector of the matrix, $A$, using equation (7.2). The basic mathematical reasoning underlying the use of the principal right eigenvector is given in Appendix III.

$$Aw = \lambda_{max} w \tag{7.2}$$

where:

$\lambda_{max}$ = *the largest eigenvalue*

$w$ = *the principal right eigenvector*

d) Test the consistency of the pairwise comparisons. Saaty (1980) developed a useful consistency measure, *CM*, defined by equation (7.3).

$$CM = \frac{CI}{R} \tag{7.3}$$

where:

$CI$ = *consistency index*

$R$ = *the computed CI of randomly generated matrices*

*CI* is computed as follows:

$$CI = \frac{\lambda_{max} - s}{s - 1} \tag{7.4}$$

where $s$ is the size of the matrix, $A$.

According to Saaty (1980), a value of $CM \leq 0.1$ is considered acceptable. Otherwise, it is necessary to reduce the inconsistencies by revising the pairwise comparisons.

e) From the principal right eigenvector, $w$, each multiplier, $M_{rf}$, is computed by using equation (7.5).

$$M_{rf} = \frac{w_{rf}}{w_{uf}} \tag{7.5}$$

where:

$w_{rf}$ = *the extracted weight corresponding to reference project r and influence factor f*

$w_{uf}$ = *the extracted weight corresponding to upcoming project u and influence factor f*

### 7.1.5 Estimation of the Upcoming Project's Productivity

Estimate the upcoming project's productivity, $P_{ur}$, using each reference project $r$.

$$P_{ur} = P_r \prod_{f=1}^{m} M_{rf} \tag{7.6}$$

where:

$P_r$ = *the productivity of reference project r*

$M_{rf}$ = *a multiplier which adjusts the productivity of reference project r due to the influence of factor f*

$m$ = *number of influencing factors*

### 7.1.6 Estimation of the Design Effort of the Upcoming Project

Estimate the design effort of the upcoming project, $E_u$, using equation (7.7).

$$E_u = \frac{\sum_{r=1}^{nr} E_{ur}}{nr} \tag{7.7}$$

where:

$E_{ur}$ = *the estimated design effort of upcoming project u in man-months using reference project r*

$nr$ = *number of reference projects*

$E_{ur}$ is computed as follows:

$$E_{ur} = \frac{O_u}{P_{ur}} \tag{7.8}$$

where:

$O_u$ = *the output of upcoming project u*

$P_{ur}$ = *the estimated productivity of upcoming project u using reference project r*

## 7.2 Application of the Model

The proposed analogy-based model was applied to the historical data obtained from NN and CMC (Appendix I). For each company, each project in the list of projects was assumed to be the upcoming project. Its design effort was then estimated using the other projects as references. It is acknowledged that using historical data may bias the results in favor of the proposed model. However, due to time limitation, comparing actual effort with the ex post estimates was the only way to test the accuracy of the model empirically.

The six steps listed in Section 7.1 and described in Sections 7.1.1-7.1.6 were applied as follows.

1) For each company, each project in the list of projects was assumed to be the upcoming project. Then, its design effort was estimated using the other projects as references.
2) The productivity of each reference project, $P_r$, was computed using equation (7.1).
3) The collected data indicated that in each company, the projects shared many common characteristics such as the design team, the method of communication, the formal design process, and automated design tools, etc. However, in addition to product complexity, *PC*, severity of requirements was identified as another major factor that had variable influence on the productivity of the projects.
4) To compute the multipliers that capture the effect of severity of requirements on design effort, a project manager from each company was asked to make pairwise comparisons using the influence rating scale shown in Table 7.1. The pairwise comparison matrices as supplied by the project managers for NN and CMC are shown in Tables 7.3 and 7.4, respectively. The computed maximum eigenvalue, $\lambda_{max}$, for each of these matrices are 5.39 and 12.82, respectively. The extreme right column of each table gives the vector of relative influence for severity requirements for each project. It is worth mentioning that the comparison values show reasonable consistency; the values of *CM* are 0.088 and 0.051 for companies 1 and 2, respectively. From the eigenvector, each multiplier, $M_{rf}$, is computed by using equation (7.5).
5) The productivity of each upcoming project, $P_{ur}$, was estimated using equation (7.6).

6) Then, the design effort of each upcoming project, $E_u$, was estimated using equation (7.7).

The computed productivity for the reference projects, $P_r$, the multipliers, $M_{rf}$, produced by the principal right eigenvector of the matrix of pairwise comparisons, the estimated productivity, $P_{ur}$, the estimated design effort for each upcoming project using each reference project, $E_{ur}$, and the estimated design effort for each upcoming project, $E_u$, for NN and CMC are listed in Appendix IV.

**Table 7.3** Project comparisons with respect to severity of requirements for NN

| Project number | 1 | 2 | 3 | 4 | 5 | Relative influence |
|---|---|---|---|---|---|---|
| 1 | 1 | 1/2 | 2 | 1 | 1/2 | 0.36 |
| 2 | 2 | 1 | 1/2 | 1/2 | 1/2 | 0.33 |
| 3 | 1/2 | 2 | 1 | 1/2 | 1/3 | 0.30 |
| 4 | 1 | 2 | 2 | 1 | 1 | 0.51 |
| 5 | 2 | 2 | 3 | 1 | 1 | 0.64 |

**Table 7.4** Project comparisons with respect to severity of requirements for CMC

| Project number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | Relative influence |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1/2 | 1 | 1/3 | 1/2 | 0.20 |
| 2 | 1/2 | 1 | 1/2 | 1 | 1/2 | 1/2 | 1/2 | 1 | 1/3 | 1/3 | 1/5 | 1/3 | 0.11 |
| 3 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1/2 | 1 | 1/3 | 1/3 | 0.19 |
| 4 | 1/2 | 1 | 1/2 | 1 | 1 | 1/2 | 1/2 | 1 | 1/3 | 1 | 1/5 | 1/3 | 0.13 |
| 5 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1/3 | 1 | 1/3 | 1/3 | 0.18 |
| 6 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 3 | 1/3 | 1 | 1/3 | 1/3 | 0.21 |
| 7 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 3 | 3 | 1 | 1/3 | 1/3 | 0.28 |
| 8 | 1 | 1 | 1 | 1 | 1 | 1/3 | 1/3 | 1 | 1/3 | 1/3 | 1/3 | 1/2 | 0.13 |
| 9 | 2 | 3 | 2 | 3 | 3 | 3 | 1/3 | 3 | 1 | 1/3 | 1/2 | 1 | 0.33 |
| 10 | 1 | 3 | 1 | 1 | 1 | 1 | 1 | 3 | 3 | 1 | 1/3 | 1/3 | 0.27 |
| 11 | 3 | 5 | 3 | 5 | 3 | 3 | 3 | 3 | 2 | 3 | 1 | 1 | 0.56 |
| 12 | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 1 | 3 | 1 | 1 | 0.48 |

## 7.3 Performance Evaluation Results

Figures 7.1 and 7.2 are scatter plots of the actual project man-months versus the analogy-based model estimated man-months for NN and CMC, respectively. As can be seen in these figures, the actual and the estimated man-months are fairly close. This observation can be further confirmed by examining the results in Tables 7.5 and 7.6. As shown in

Table 7.5, the computed *MMRE* for NN data is 13% and four of the five projects had an error rate less than or equal to 20%. As shown in Table 7.6, the computed *MMRE* for CMC data is 14%, and 83% of the model estimates are within 23% of the actual values. Thus, it can be concluded that the proposed model is reasonably accurate.

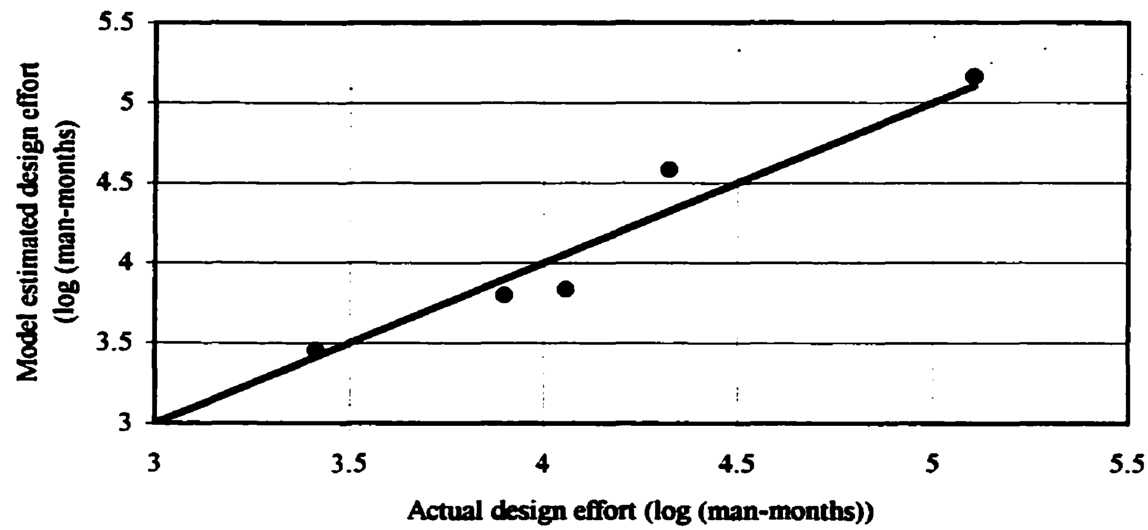



**Figure 7.1** Actual design effort versus analogy-based model estimates for NN

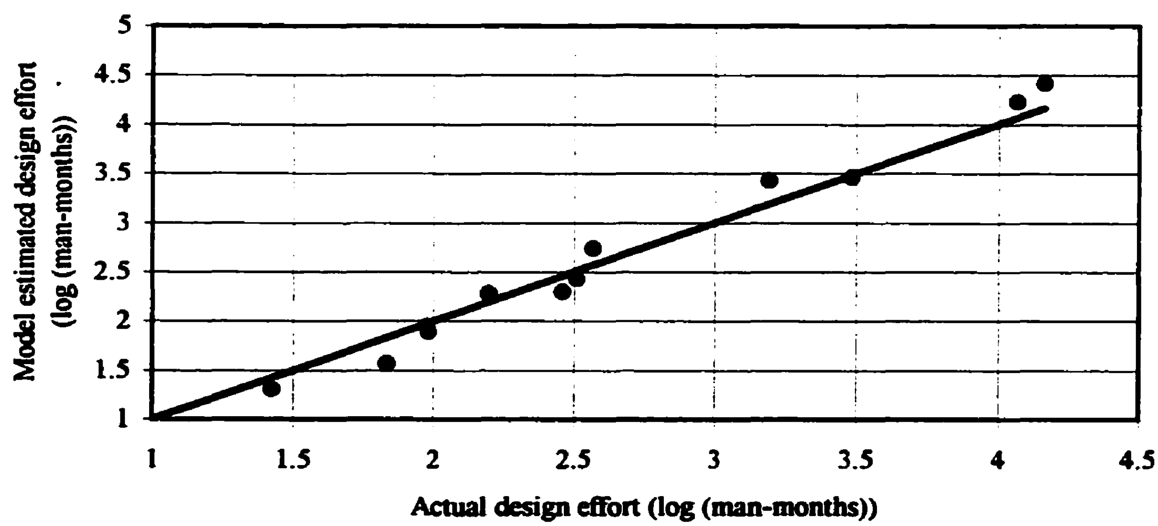



**Figure 7.2** Actual design effort versus analogy-based model estimates for CMC

**Table 7.5** Analogy-based model evaluation results for NN

| Upcoming project | Actual design effort (man-months) | Estimated design effort (man-months) | Error (%) |
|---|---|---|---|
| 1 | 30.37 | 31.39 | 3 |
| 2 | 57.89 | 46.21 | 20 |
| 3 | 49.34 | 45.11 | 9 |
| 4 | 75.45 | 96.29 | 28 |
| 5 | 164.69 | 176.98 | 7 |
| | | | *MMRE* = 13 |

**Table 7.6** Analogy-based model evaluation results for CMC

| Upcoming project | Actual design effort (man-months) | Estimated design effort (man-months) | Error (%) |
|---|---|---|---|
| 1 | 6.26 | 4.81 | 23 |
| 2 | 4.16 | 3.71 | 11 |
| 3 | 7.26 | 6.63 | 9 |
| 4 | 7.23 | 7.04 | 3 |
| 5 | 8.99 | 9.77 | 9 |
| 6 | 12.33 | 11.33 | 8 |
| 7 | 13.06 | 15.47 | 18 |
| 8 | 11.69 | 9.97 | 15 |
| 9 | 32.57 | 31.96 | 2 |
| 10 | 24.35 | 30.90 | 27 |
| 11 | 58.44 | 68.99 | 18 |
| 12 | 64.44 | 82.95 | 29 |
| | | | *MMRE* = 14 |

## 7.4 Summary

As a complementary method to the techniques described in the previous two chapters, this chapter proposed an analogy-based model for estimating design effort. The model can be compactly expressed as:

$$E_u = \frac{1}{nr}\sum_{r=1}^{nr}\frac{O_u}{P_r\prod_{f=1}^{m}M_{rf}} \tag{7.9}$$

where:

$E_u$ = *the estimated design effort of upcoming project u in man-months*

$O_u$ = *the output of upcoming project u*

$P_r$ = *the productivity of reference project r*

$M_{rf}$ = *a multiplier which adjusts the productivity of reference project r due to the influence of factor f*

$m$ = *number of influencing factors*

$nr$ = *number of reference projects*

The model is based on a newly defined productivity ratio and on estimating the influence of significant factors on productivity using the eigenvector approach. The model is intuitive and does not require any prior relationships to be developed. The application of the model to the data from NN and CMC indicates that the model is reasonably accurate.

It is worth mentioning that while the new productivity ratio was introduced for estimating design effort, it is also useful for comparing the performance between various projects and in helping identify the possible causes for decreases or increases in productivity.

The level of error in the analogy-based model for the NN and CMC cases was about the same as that found for the neural network and parametric models.

# Chapter 8

# Applications of the Models

The previous three chapters described a number of new models for estimating design effort based on product complexity. However, in addition to estimating quantitatively the effort required to design a product, these models can be useful in several other ways. They can be used to estimate cost, staffing patterns, and project duration. Most importantly, they can be used to study how these parameters will be affected due to changes in design requirements, resource allocation, etc. (Bashir and Thomson, 1999f, 1999g). This chapter demonstrates these applications through demonstrative examples.

## 8.1 Project Cost Estimation

An important aspect of any design project is to estimate how much it will cost. Multiplication of estimated design effort in hours, $\hat{E}$, using one or a combination of the models described in the previous chapters, times the average cost per hour, $AC$, gives an estimate for total direct manpower cost, $T$. Since labor costs make up the majority of the cost for most development environments, $T$ provides a good estimate of project cost.

$$T = \hat{E}\,AC \tag{8.1}$$

Equation (8.1) can be more complicated if the different cost structures for managers, designers, etc., are included.

## 8.2 Project Duration Estimation

The main application of Norden's model ($y' = 2\hat{E}\alpha t e^{-\alpha t^2}$ (2.1)) is to estimate the required design effort at any point of time. In addition, it can be used to estimate project duration. However, to do so, the required design effort, $\hat{E}$, and the shape parameter, $\alpha$, of Norden's effort-time model need to be estimated. The required design effort can be estimated by using one or a combination of the models described in the previous three chapters. The shape parameter can be estimated subjectively by estimating the peak design effort or the time to reach it. It is worth mentioning that to minimize the use of pure subjectivity, Putnam (1978a) found an interesting empirical relationship between effort, time to peak, and the difficulty of a software project. Such a relationship needs to be developed for design projects in general. However, this needs a large number of historical projects to be analyzed and was not possible in this thesis. Alternatively, the average ratio of design effort to time-to-peak, $\gamma$, for a number of previously completed projects can be used to estimate project duration (Bashir and Thomson, 1999f, 1999g).

Subsistuting $t_o = \dfrac{\hat{E}}{\gamma}$ in equation (2.2), $\alpha = \dfrac{0.5}{t_o^2}$, yields:

$$\alpha = \frac{0.5\gamma^2}{\hat{E}^2} \tag{8.2}$$

where:

$\hat{E}$ = *estimated design effort in appropriate units, e.g. hours or man-months*

$\gamma$ = *the average ratio of actual design effort to time-to-peak for a number of previously completed projects in appropriate units, e.g. man-months per month or hours per month. The selected units should be compatible to that of* $\hat{E}$.

Since the curve obtained by Norden's model tails out to infinity, a method is needed to estimate a project's end point. The average ratio of actual design effort spent during the last month to the actual total design effort, $x$, for a number of previously completed projects is used to estimate project duration, $t_d$, as follows.

From equation (2.3), $y = \hat{E}(1 - e^{-\alpha t^2})$, the cumulative manpower up to $t_d - 1$ is given by:

$$\hat{E}(1 - x) = \hat{E}(1 - e^{-\alpha(t_d - 1)^2}) \tag{8.3}$$

Rearranging equation (8.3) gives:

$$t_d = \sqrt{\frac{-\ln x}{\alpha}} + 1 \tag{8.4}$$

where:

$t_d$ = *project duration in months*

From equations (8.2) and (8.4)

$$t_d = \sqrt{\frac{-2\hat{E}^2 \ln x}{\gamma^2}} + 1 \tag{8.5}$$

The relationship between product complexity, $PC$, and project duration, $t_d$, can be determined by substituting $\hat{E} = aPC^b$ into equation (8.5).

$$t_d = \sqrt{\frac{-2(aPC^b)^2 \ln x}{\gamma^2}} + 1 \tag{8.6}$$

Thus, the method for estimating project duration can be summerized as follows:

- Estimate the required design effort, $\hat{E}$, using one or a combination of the models described in Chapters 5, 6, and 7.
- Compute the average ratio of actual design effort to time-to-peak, $\gamma$, and the average ratio of actual design effort spent during the last month to the actual total design effort, $x$, for a number of previously completed projects.
- Estimate project duration by substituting for $\hat{E}$, $\gamma$, and $x$ values in equation (8.5).

**Example 8.1**

Given the model for Nortel Networks, $\hat{E} = 24.33PC^{1.39}$ (5.4), $x = 0.04$, and $\gamma = 1728$ hours per month, a useful chart is obtained by substituting different values of $PC$ into equations (5.4) and (8.6). The resulting graph shown in Figure 8.1, gives the relationship between product complexity (horizontal axis), design effort in man-months[15] (left-hand axis), and duration in months (right-hand axis). The chart can be used to obtain quick estimates for design effort and duration at different levels of product complexity. For example, a project will require 18 man-months of design effort and will need 5 months to be completed if its corresponding product complexity is 30.

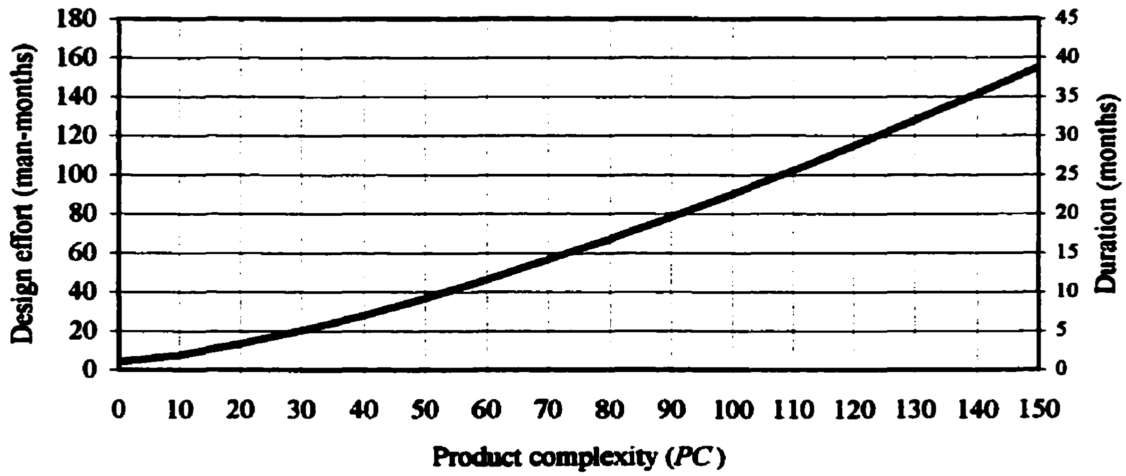



**Figure 8.1** The relationship among product complexity, design effort, and duration based on a set of hypothetical design projects

## 8.3 Change in Design Requirements

Design requirements are demands and wishes that clarify the design task in the space of needs (Pahl and Beitz, 1984). These requirements are usually set at the beginning of the project. However, while development is in progress, many changes can arise. These changes can lead to a change in product complexity, severity of requirements, or other parameter, and thus, a change in the estimate of project effort and duration. In addition, these changes can lead to redoing work already done.

For a change in design requirements which leads to an increase in product complexity, a relationship between initial product complexity, $PC_i$, new product complexity, $PC_n$, and a

[15] Man-months were obtained by dividing the number of hours by 152.

new estimate for project duration, $t_{dn}$, can be approximated. A new estimate for project effort, $\hat{E}_n$, is given by:

$$\hat{E}_n = aPC_n^b - (aPC_i^b(1-e^{-\frac{0.5t^2\gamma^2}{(aPC_i^b)^2}}) - aPC_n^b(1-e^{-\frac{0.5t^2\gamma^2}{(aPC_n^b)^2}})) + RW \tag{8.7}$$

where:

*RW = time to be spent on rework in man-months*

Equation (8.7) can be rewritten as:

$$\hat{E}_n = 2aPC_n^b - aPC_n^b e^{-\frac{0.5t^2\gamma^2}{(aPC_n^b)^2}} - aPC_i^b + aPC_i^b e^{-\frac{0.5t^2\gamma^2}{(aPC_i^b)^2}} + RW \tag{8.8}$$

Substituting $\hat{E}_n$ in equation (8.5) gives:

$$t_d = \sqrt{\frac{-2(2aPC_n^b - aPC_n^b e^{-\frac{0.5t^2\gamma^2}{(aPC_n^b)^2}} - aPC_i^b + aPC_i^b e^{-\frac{0.5t^2\gamma^2}{(aPC_i^b)^2}} + RW)^2 \ln x}{\gamma^2} + 1} \tag{8.9}$$

**Example 8.2**

Again, using the model for Nortel Networks as in Example 8.1 and substituting $a$ = 24.33, $b$ = 1.39, $\gamma$ = 1728 hours per month, and $x$ = 0.04, and $PC_i$ = 90 in equation (8.6) gives an initial project duration estimate, $t_d$, of 20 months. If, after 3 months ($t$ = 3), some changes in design requirements have led to a new estimate for product complexity, ($PC_n$ = 120) and no rework is required ($RW$ = 0), then, substituting for $a$, $b$, $PC_i$, $PC_n$, $\gamma$, $t$, and $x$ in equation (8.9) gives a new estimate for project duration, $t_{dn}$, of 28 months. Based on this, the expected time slippage is about 8 months.

Both schedules are shown graphically in Figure 8.2 where the solid curve represents the initial schedule, and the dashed curve represents the new schedule. One interesting fact supported by Figure 8.2 is that projects that have many changes in requirements at different points of time have noisy manpower utilization data.

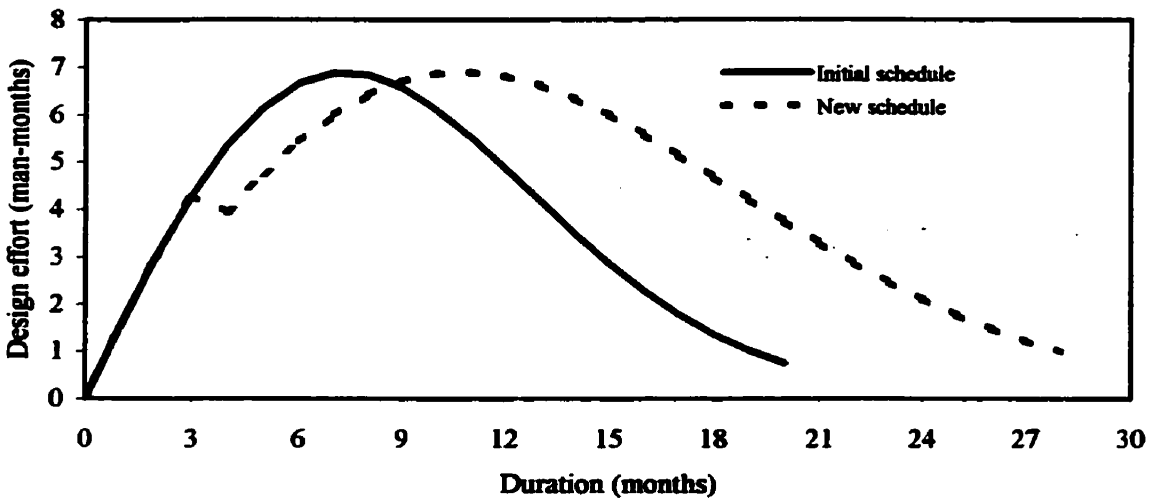



**Figure 8.2** Impact on project duration for a change in product complexity from 90 to 120 after 3 months of project activity

## 8.4 Change in Resource Allocation

Using Norden's model, it is simple to derive a family of curves of design effort versus project duration for different resource levels. This is shown in Figure 8.3. If at a certain point of time while development is in progress, it is decided to shorten project duration by allocating more resources, either by extending working hours or adding more people, decision-makers need to estimate how many resources must be added in order to reduce project duration. This can be done by substituting in equation (2.1) a new estimate for the shape parameter, $\alpha_n$, that gives the desired new project duration.

$\alpha_n$ is simply obtained by rearranging equation (8.4),

$$\alpha_n = \frac{-\ln x}{(t_d - 1)^2} \tag{8.10}$$

where:

*$t_d$ = the desired project duration in months*

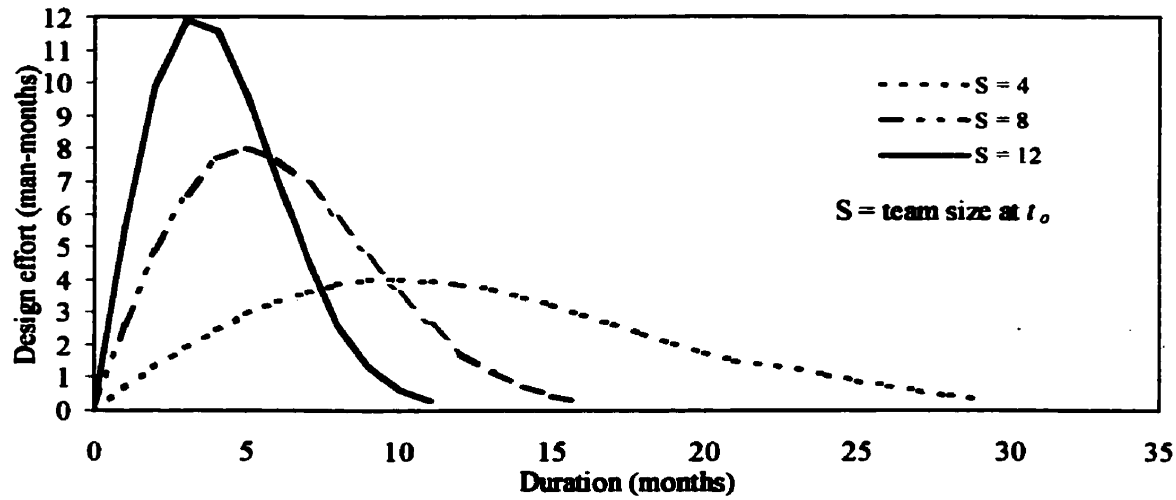



**Figure 8.3** Impact of resource levels on project duration for a hypothetical project

__Example 8.3__

Consider the initial schedule of the project in Example 8.2. If after 3 months ($t$ = 3), it is decided to shorten the project duration by 4 months, then, substituting for $x$ = 0.04 and $t_d$ = 16 months in equation (8.10) gives a new estimate for shape parameter, $\alpha_n$, of 0.014. The use of Norden's model (equation (2.1)) gives the required monthly design effort needed for the project as shown in Figure 8.4. Note that area *ghi* has been added to the initial project effort estimate. This is an artifact of the estimation technique since this time has actually passed. The increase in effort represented by the area *ghi* will have to be compensated for in the total effort of the project, although it can be disregarded if small enough.

Obviously, as more resources are added, project duration decreases. However, this does not imply that more resources can be added indefinitely. In fact, adding more resources to a project is not always useful. Rather, it depends on the availability of activities that can be tackled by the added resources. Thus, using equation (8.10) improperly can give false results. In fact, adding personnel above a certain level can cause project duration to increase.

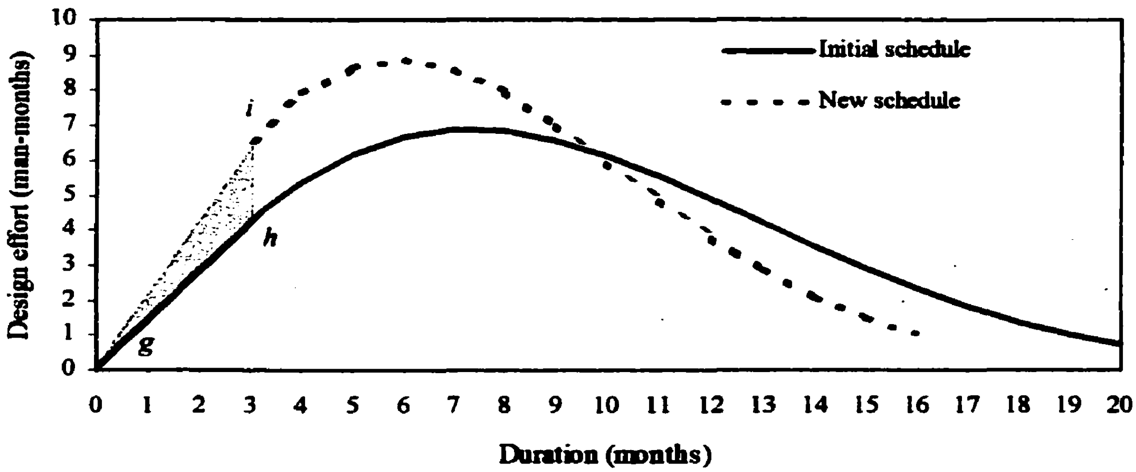



**Figure 8.4** Impact of changes in resource allocation for a change in project duration from 20 to 16 months

## 8.5 Summary

Through demonstrative examples, this chapter presented some of the applications of the effort estimation models. As summarized in Figure 8.5, these applications constitute a complete quantitative, estimation methodology that provides initial as well as updated project estimates from feasibility study to project completion. The complete methodology provides answers to the following vital questions:

- How much design effort is needed?
- How much will the project cost?
- How much manpower is needed at any given time?
- How long will the project take?
- What will happen if changes in design requirements or resource allocation are carried out at a certain point of time while development is in progress?
- Is the project feasible?

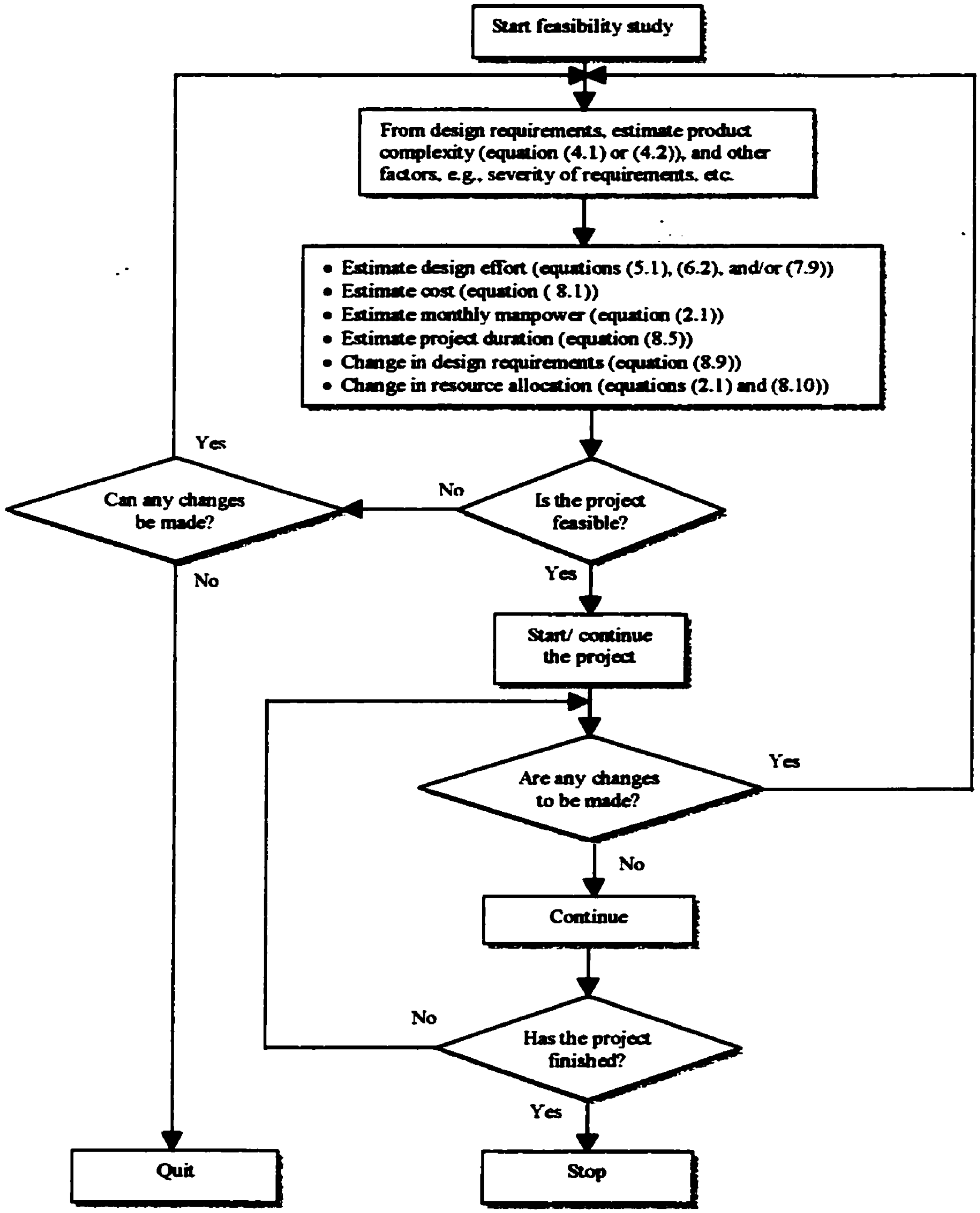



**Figure 8.5** Logical outline of the estimation methodology

# Chapter 9

# Conclusions

*"No single work advances understanding very far. The aims of a scientific work are limited by the formal character of the theory, by the phenomena it encompasses, by experimental situations it uses, by the types of subjects it studies, and the data it gathers. Of course a theory may speak beyond its initial base--all scientists hope for just that. But science is a series of successive approximations. Not all things can be done at once, and if one aspires to go far, he must start somewhere. If one aims at covering of all human thinking in a single work, the work will necessarily be superficial. If one aims at probing in depth, then many aspects of the subject, however important, will be untouched"*

(*Newell and Simon 1972*)

The preceding chapters have described new models for estimating design effort, and have demonstrated some of their applications. This chapter begins with a summary of the findings; then, it proceeds to identify the major limitations of this research. The chapter concludes with several suggested areas for further research.

## 9.1 Summary of Findings

The literature review revealed that one of the major factors contributing to the problem of overruns is the inadequacy of available effort estimation techniques. It also revealed that domains where complexity metrics are employed such as software, are more successful at effort estimation than those that do not. Thus, a potentially important step in effort estimation improvement is to explicitly incorporate in the estimation model a metric that realistically reflects the complexity of the product to be designed.

A new metric, *PC*, which is based on measuring the complexity of the functional tree of a product was presented. Using the Kruskal-Wallis test for analyzing data obtained from an experiment confirmed the validity of the metric as a predictor of design effort. These results were confirmed further by analyzing data from historical projects from Nortel Networks (NN), Canadian Marconi Company (CMC), and General Electric Hydro (GE) using the Spearman rank-order correlation coefficient.

Using traditional regression analysis models for NN, CMC, and GE were developed. All the models were based on the new product complexity metric, *PC*. Analyzing the performance of the developed models indicated generally that the models were reasonably accurate.

The use of general models to estimate the design effort for different companies was tested by developing models from the combined data from Nortel and CMC. However, the performance of these models was unsatisfactory. This is probably due to the heterogeneity of the products and/or to fact that the differences in the different environments of the companies are not reflected in just two parameters.

One of the major disadvantages of using traditional regression analysis is the need to assume the form of the regression function. If the form is incorrectly chosen, then, the regression results in a poor fit with the data, and consequently, poor estimations. Thus, there is a need for a method that does not depend on a presumed function. Artificial neural networks is such a tool that meets this requirement. Using the data sets from CMC and GE, General Regression Neural Network (GRNN) models were designed, trained, and tested to estimate design effort. Within the limited data sets, the developed neural network model produced good results, and were comparable to those from regression analysis.

One of the shortcomings of parametric and artificial neural network models is that they cannot be constructed unless there are patterns of relationships between effort as the dependent variable and one or more factors as the independent variables. In addition, in a relatively constant environment. Thus, there is a need for a method that gives the estimator a means of prediction in unusual situations and/or when parametric artificial neural network models are difficult to obtain for estimation. For this purpose, an analogy-based model for estimating design effort was proposed. The model incorporates an eigenvector approach to estimate the influence of the various factors on productivity. In addition to its potential to derive better estimates, the eigenvector approach provides a measure of consistency, which is not available in a direct estimation method. The model was applied to the data from historical projects from CMC and GE. Analyzing the performance of the model indicated that the model was comparable in performance to the parametric and artificial neural network models.

The models were developed with several goals in mind:

- applicable to a wide range of engineering projects,
- reasonably accurate,
- easy to use, and
- parsimonious.

The models have met all of these goals to some extent. They are based on functionality; therefore, they have the potential for being more widely applicable than many others. According to a number of estimation accuracy tests, they performed well. They are relatively easy to use, and in addition, they use a small number of inputs.

In terms of estimation accuracy, the results indicated that the models are comparable and significantly better than the companies' original estimations. One general important conclusion that can be drawn from these results is that the use of functionality to estimate design effort is more accurate than current methods.

In terms of their characteristics, Table 9.1 indicates that each model has its own characteristics that make it unique, but none is better than any other in regard to all aspects.

**Table 9.1** A general comparison between the parametric model, the GRNN model, and the analogy based model

| Aspect | Parametric model | GRNN model | Analogy-based model |
|---|---|---|---|
| **Prior assumptions about the equation's form** | The equation's form needs to be known or guessed | No equation's form needs to be known or guessed. However, some trials are required to set up the model's parameters | No equation's form needs to be known or guessed |
| **Handling more than one output simultaneously** | Only one output | It can handle many outputs simultaneously | Only one output |
| **Performing sensitivity analysis** | Sensitivity analysis can be easily performed | Sensitivity analysis can be performed with some difficulty | Sensitivity analysis can be performed with some difficulty |
| **Computations** | Some computations are required | Tremendous computations are required. It cannot be constructed without using a computer | Some computations are required |
| **Model accessibility** | Accessible | Not accessible | Accessible |
| **Dependency on the experience of the estimator** | Low | Low | More than the parametric and GRNN models, but less than pure expert judgements |
| **The use in unusual situations** | It may give inaccurate results | It may give inaccurate results | More reliable than the parametric and GRNN models |

Finally, the applications of the models described in Chapter 8 showed that a complete quantitative estimation methodology that addresses many vital questions is now available. These questions include:

- How much will the project cost?
- How much manpower is needed at any time?
- How long will the project take?
- What will happen if changes in design requirements or resource allocation are carried out at a certain point of time while development is in progress?

## 9.2 Limitations of this Research

1) The product complexity metric, *PC*, cannot be computed without a thorough understanding of the required functionality of the product to be designed. Furthermore, because functional decomposition involves a degree of subjectivity, the suggested guidelines and the decomposition steps still do not guarantee that different users will obtain 100% identical product decompositions and product complexity estimations. However, demanding that any useful metric should measure an attribute without any possible variation or error would rule out the use of many useful metrics. Since there is little need to have global consistency for practical applications, consistency for the product complexity metric is only required within an organization. Therefore, what is important is that all the products within a company are decomposed consistently. This can be achieved by making one analyst or a group of analysts responsible for all decomposition tasks and/or by defining standard procedures for creating hierarchies of functions.
2) It also recognized that in spite of all the precautions that have been made, the experiment used to determine if the complexity metric, *PC*, was a good predictor of design effort, described in Chapter 4, had some weaknesses and limitations, such as:
   - Because of cost and time limitations the tasks were well-defined problems.
   - Only one type of problem was involved.

- The sample size used for the analysis was small. This means that even though subjects were randomly assigned to the tasks, it cannot be guaranteed that the experimental outcome was not affected by differences in skill among the subjects.

3) It is acknowledged that only using historical data may bias the results in favor of the estimation models developed in this thesis. Due to time limitation, comparing actual effort with the ex post estimates was the only way to test the accuracy of the models empirically.
4) Due to the limitations in time and data availability, the models for estimating project duration were not tested.

## 9.3 Recommendations for Further Research

Models to estimate design effort hold promise, not only in providing accurate estimates, but also in aiding the understanding of those factors that have the largest potential for improving productivity. As more is learned about the design process, better models of the relationships among the factors involved can be built. While some progress has been made in this research, a great deal remains to be done. Some suggestions are given below.

1) The developed product complexity metric, *PC*, has been shown to be a good predictor of design effort. Nevertheless, other metrics may be possible and should be explored.
2) More controlled experiments using larger sample sizes are definitely needed to test the use of the product complexity metric, *PC*. In addition, factors other than product complexity should be investigated.
3) If there is a high similarity of conditions among the projects under consideration, the candidate factors to be included in the analogy-based model can be easily determined. However, as heterogeneity increases, identification of such factors becomes more difficult. It is suggested that this difficulty can be overcome by using the Analytic Hierarchy Process. Research is needed to confirm this suggestion.
4) Putnam (1978a) found an interesting empirical relationship between effort, $E$, time to peak, $t_o$, and the difficulty of a software project. Further research is needed to explore such a relationship for design projects in general.

5) This dissertation provided one test of the usefulness of ANNs in estimating design effort. It should be noted that the goal was not to seek additional research uses for artificial neural networks. Rather it was to find a more accurate and practical estimation tool. However, the capabilities of ANNs in estimating design effort need to be further investigated in cases where mathematical relationships cannot be easily established.

# References


Adrangi, B., W. Harrison, 1987, Effort Estimation in a System Development Project, *Journal of Systems Management*, Vol. 38, No. 8, pp. 21-23.

Albrecht, A. J., 1979, Measuring Application Development Productivity, *Proceedings of Joint Share, Guide, and IBM Application Development Symposium*, pp. 83-92.

Bahill, A. T., W. L. Chapman, 1995, Case Studies in System Design, *International Symposium and Workshop on Systems Engineering of Computer Based Systems*, Piscataway, New Jersey, pp. 43-50.

Bailey, D., D. Thompson, 1990, How to Develop Neural-network Applications, *AI Expert*, June, pp. 38-47.

Bashir, H. A., V. Thomson, 1999a, Metrics for Design Projects: a Review, *Design Studies*, Vol. 20, No. 3, pp. 263-277.

Bashir, H. A., V. Thomson, 1999b, Estimating Design Complexity, *Journal of Engineering Design*, Vol. 10, No. 3., pp.247-257

Bashir, H. A., V. Thomson, 1999c, Models for Estimating Design Effort and Time, *Design Studies*, (in press).

Bashir, H.A., V. Thomson, 1999d, An Analogy-based Method for Estimating Design Effort, *Design Studies*, (in press).

Bashir, H.A., V. Thomson, 1999e, Estimating Design Effort Using Artificial Neural, Networks, *Proceedings of the 3rd International Conference on Engineering Design and Automation*, Vancouver, Canada.

Bashir, H. A., V. Thomson, 1999f, A Quantitative Estimation Methodology for Design Projects, *Proceedings of the International Conference on Industrial Engineering and Production Management (IEPM'99)*, Glasgow, U.K., pp. 498-506.

Blessing, L. T. M., 1994, *A Process-Based Approach to Computer-Supported Engineering Design*, Ph.D. Dissertation, University of Twente, Enschede, The Netherlands.

Boehm, B. W., 1981, *Software Engineering Economy*, Prentice-Hall, Englewood Cliffs, New Jersey.

Bounds, G., 1998, The Last Word on Project Management, *IIE Solutions*, November, pp. 41-43.

Bronikowski, R. J., 1986, *Managing the Engineering Design Function*, Van Nostrand Company, New York.

Brooks, F. P., 1975, *The Mythical Man-month: Essays on Software Engineering* Addison-Wesley Publishing Company, Don Mills, Ontario.

Cacoullos, T., 1966, Estimation of a Multivariate Density, *Annals of the Institute of Statistical Mathematics*, Vol. 18, No. 2, pp. 179-189.

Carpenter, G., S. Grossberg, 1987, ART-2: Self-Organization of Stable Category Recognition Codes for Analog Input Patterns, *Applied Optics*, Vol. 26, No. 23, pp. 4919-4930.

Clark, K. B., W. B. Chew, T. Fujimoto, 1987, Product Development in the World Auto Industry, *Brookings Papers on Economic Activity*, Vol. 3, pp. 729-771.

Clark, K. B., 1989, Project Scope and Performance: the Effect of Parts Strategy and Supplier Involvement on Product Development, *Management Science*, Vol. 35, No. 10, pp. 1247-1263.

Conte, S. D., H. E. Dunsmore, V. Y. Shen, 1986, *Software Engineering Metrics and Models*, The Benjamin /Cummings Publishing Company, Menlo Park.

Cook, M. L., 1982, Software Metrics: an Introduction and Annotated Bibliography, *Software Engineering Notes (ACM SIGSOFT)*, Vol. 7, No. 2, pp.41-60

Cooper, R. G., 1990, Stage-gate System: a New Tool for Managing New Products, *Business Review Horizons*, May-June, pp. 44-54.

Cordero, R., 1991, Managing for Speed To Avoid Product Obsolescence: A Survey of Techniques, *Journal of Product Innovation Management*, Vol. 8, pp. 283-294.

Cote, V., S. Bourque, S. Oligny, N. Rivard, 1988, Software Metrics: an Overview of Recent Results, *The Journal of Systems and Software*, Vol. 8, pp. 121-131.

Cross, N., 1994, *Engineering Design Methods*, John Wiley & Sons, New York.

DeMarco, T., 1982, *Controlling Software Projects*, Yourdon Press, New York.

Dreger, J. B., 1989, *Function Point Analysis*, Prentice-Hall, Englewood Cliffs, New

Evans, S., 1990, Implementing Framework for Integrated Design Teams, *Journal of Engineering Design*, Vol. 1, No 4, pp. 355-363.

Elmaghraby, S. E., W. S. Herroelen, 1980, On the Measurement of Complexity in Activity Networks, *European Journal of Operational Research*, Vol. 5, pp. 223-234.

Elsayed, E. A., T. O. Boucher, 1985, *Analysis and Control of Production Systems*, Prentice-Hall, Inc., Englewood Cliffs.

Eyman, R. K., C. E. Meyers, R. Bendel, 1973, New Methods for Test Selection and Reliability Assessment Using Stepwise Multiple Regression and Jackknifing, *Educational and Psychological Measurement*, Vol. 33, pp. 883-894.

Fenton, N. E., S. L. Pfleeger, 1997, *Software Metrics: A Rigorous and Practical Approach*, PWS Publishing Company, Boston.

Gioia, J., 1996, Twelve Reasons Why Programs Fail, *PM Network*, Novemeber, pp. 16-19

Gomory, R. E., 1989, From the Ladder of Science to the Product Development Cycle, *Harvard Business Review*, Vol. 68, No. 6, pp. 99-105.

Gott, B., 1980, Why Can't We Manage CAD? *Proceedings of conference on managing Computer Aided Design*, Institution of Mechanical Engineers, London.

Griffin, A., 1993, Metrics for Measuring Product Development Cycle Time, *Journal of Product Innovation Management*, Vol. 10, No. 2, pp. 112-125

Griffin, A., 1997, The Effect of Project and Process Characteristics on Product Development Cycle Time, *Journal of Marketing Research*, Vol. 34, No. 1, pp. 24-35.

Hajek, V. G., 1984, *Management of Engineering Projects*, McGraw-Hill Book Company, New York.

Hales, C., 1987, *Analysis of the Engineering Design Process in an Industrial Context*, Ph.D. Dissertation, University of Cambridge.

Hecht-Nielsen, R., 1988, Applications of Counterpropagation Networks, *Neural Networks*, Vol. 1, No. 2, pp. 131-139.

Heesterman, A. R. G., 1990, *Matrices and their Roots*, World Scientific Publishing Company, Singapore.

Helmer, O., 1966, *Social Technology*, Basic Books.

Hinton, G. E., T. J. Sejnowski, 1986, Learning and Relearning in Boltzmann Machines. In *Parallel Distributed Processing*, MIT Press, Cambridge, Mass, pp. 282-317.

Hopfield, J. J., 1982, Neural Networks and Physical Systems with Emergent Collective Computational Abilities, *Proceedings, National Academy of Science*, Vol. 79, pp. 2554-2558.

Hubka, E., 1982, *Principles of Engineering Design*, Butterworth & Co (Publishers) Ltd., London.

Hubka, E., 1988, *Theory of Technical Systems: a Total Concept Theory for Engineering Design*, Springer-Verlag, New York.

Hughes, R. T., 1996, Expert Judgement as an Estimating Methods, *Information and Software Technology*, Vol. 38, No 2, pp. 67-75.

Jacome M. F., V. Lapinskii, 1997, NREC: Risk Assessment and Planning of Complex Designs, *IEEE Design & Test of Computers*, Vol. 14, No. 1, pp. 42-49.

Jeffery, D. R., 1987, The Relationship Between Team Size, Experience, and Attitudes and Software Development Productivity, *COMPSAC*, IEEE Computer Society Press, pp. 2-8.

Jones, C., 1986, *Programming Productivity*, McGraw-Hill Book Company.

Kannapan, S. M., 1995, Function Metrics for Engineered Devices, *Applied Artificial Intelligence*, Vol. 9, No. 1, pp. 45-64.

Kelley, J. E. Jr., M. R. Walker, 1959, Critical-Path Planning and Scheduling, *Proceedings of Eastern Joint Computer Conference*, Boston, pp. 160-173.

Kingel, A. R., 1966, Bias in PERT Project Completion Time Calculations for a Real Network, *Management Science*, Vol. 13, No. 4, pp.194-201.

Kosko, B., 1987, Bi-directional Associative Memory, *IEEE on Systems, Man and Cybernetics*, Vol. 18, No. 1, pp. 49-60.

Kota, S., A. C. Ward, 1990, Functions, Structures, and Constraints in Conceptual Design, *Proceedings of the 2nd International Conference on Design Theory and Methodology*, Chicago, Illinois, pp.239-250.

Krishnamoorthy, C. S., S. Rajeev, 1996, *Artificial Intelligence and Expert Systems for Engineers*, Boca Raton: CRC Press.

Kusiak, A., E. Szczerbicki, 1992, A Formal Approach to Specifications in Conceptual Design, *Journal of Mechanical Design*, Vol. 114, pp. 659-666.

Leech, D. J., 1972, *Management of Engineering Design*, John Wiley & Sons, Belgium.

Levin, R., C. A. Kirkpatrick, 1966, *Planning and Control with PERT-CPM*, McGraw-Hill Book Company, New York.

Mansuy, J., 1991, Work Breakdown Structure: A Simple Tool for Complex Jobs, *Cost Engineering*, Vol. 33, No. 12, pp. 15-18.

Miles, L. D., 1961, *Techniques of Value Analysis and Engineering*, New York, McGraw-Hill Book Company.

Millson, M .R., S. P. Raj, D. Wilemon, 1992, A Survey of Major Approaches for Accelerating New Product Development, *Journal of Product Innovation Management*, Vol. 9, No. 1, pp. 53-69.

Moller, K. H., D. J. Paulish, 1993, *Software Metrics: a Practitioner's Guide to Improved Product Development*, Chapman & Hall Computing, London.

Mosteller, F., J. W. Tukey, 1977, *Data Analysis and Regression*, Addison-Wesley Publishing Company, Don Mills, Ontario.

Montgomery, D. C., 1984, *Design and Analysis of Experiments*, John Wiley, New York.

Murmann, P. A., 1994, Expected Development Time Reductions in the German Mechanical Engineering Industry, *Journal of Product Innovation Management*, Vol. 11, No. 3, pp. 236-252.

Newell, A., H. A. Simon, 1972, *Human Problem Solving*, Prentice Hall, Englewood Cliffs, New Jersey.

Norden, P. V., 1964, *Manpower Utilization Patterns in Research And Development Projects*, Ph.D. Dissertation, Columbia University, New York.

Norden, P. V., 1970, Useful Tools for Project Management, In *Management of Production*, ed. M. K. Starr, Penguin Books, Baltimore, MD, pp. 71-101.

Norris, K. P., 1971, The Accuracy of Project Cost and Duration Estimates in Industrial R&D, *R&D Management*, Vol. 2, No. 1, pp. 25-36.

O'Shaughnessy, K., R. Sturges, 1992, A Systematic Approach to Conceptual Engineering Design, *Design Theory and Methodology*, Vol. 42, pp. 283-291.

Pahl, G. W. Beitz, 1984, *Engineering Design: A Systematic Approach*, Design Council, London.

Parzen, E., 1962, On Estimation of a Probabilistic Density Function and Mode, *Annals of Mathematical Statistics*, Vol. 33, pp. 1065-1076.

Phan, D., D. Vogel, J. Nunamaker, 1988, The Search for Perfect Project Management, *Computerworld*, Vol.22, No. 39, pp. 95-100.

Putnam, L. H., 1978a, General Empirical Solution to the Macro Software Sizing and Estimating Problem, *IEEE Transactions on Software Engineering*, Vol. SE-4, No 4, pp. 345-361.

Putnam, L. H., 1978b, Progress in Modeling the Software Life Cycle in a Phenomenological Way to Obtain Engineering Quality Estimates and Dynamic Control of the Process, *Second Software Life Cycle Management Workshop*, Atlanta, Georgia pp. 105-128.

Rosenblatt, F., 1961, *Principles of Neorodynamics: Perceptrons and The Theory of Brain Mechanisms*, Spartan Books, Washington D.C.

Ross, B. H., 1986, Remindings in Learning: Objects and Tools, In *Similarity, Analogy, and Thought*, eds. S. Vosniadou, A. Ortony, Cambridge University Press, New York.

Rumelhart, D. E., G. E. Hinton, R. J. Williams, 1986, Learning Internal Representations by Error Propagation, In *Parallel Distributed Processing: Explorations in the Microstructures of Cognition*, eds. D. E. Rumelhart, J. L. McClelland, MIT Press, Cambridge, Mass., pp. 318-362.

Rummer, G., A. Brache, 1990, *Improving Performance: How to Manage the White Space on the Organization Chart*, Jossey-Bass, San Francisco.

Saaty, T. L., 1980, *The Analytic Hierarchy Process*, McGraw-Hill, Inc., New York.

Shillito, M. L., D. J. De Marle, 1992, *Value: Its Measurement, Design, and Management*, John Wiley & Sons, New York.

Siegel, S., N. J. Castellan, 1988, *Nonparametric Statistics for the Behavioral Sciences*, McGraw-Hill Book Company, New York.

Specht, D .F., 1988, Probabilistic Neural Networks for Classification, Mapping, or Associative Memory, *Proceedings of International Conference on Neural Networks*, Vol. 1, pp. 525-532.

Specht, D. F., 1991, A General Regression Neural Network, *IEEE Transactions on Neural Networks*, Vol. 2, No. 6, pp. 569-576.

Suh, N. P., 1990, *The Principles of Design*, Oxford University Press, Inc., New York.

Thamhain, H. J., D. L. Wilemon, 1986,Criteria for Controlling Projects According to Plan, *Project Management Journal* Vol. 17, No. 2, pp 75-81.

Ulrich, K., D. Satorius, S. Pearson, M. Jakiela, 1993, Including the Value of Time in Design for Manufacturing Decision Making, *Management Science*, Vol. 39, No. 4, pp. 429-447.

Waldron, M. B., K. J. Waldron, 1996, The influence of the Designer's Expertise on the Design Process, In *Mechanical design*, eds. M. B. Waldron, K. J. Waldron, Springer-Verlag Inc., New York.

Wallace, K. M., C. Hales, 1987, Detailed Analysis Of an Engineering Design Project, *Proceedings of the 1987 International Conference on Engineering Design*, ed. W. E. Eder, ASME, New York, Vol. 1, pp. 94-101.

Walston, C. E., C. P. Felix, 1977, A Method of Programming Measurement and Estimation, *IBM System Journal*, Vol. 16, No. 1, pp. 54-73.

Walverton, R. W., 1974, The Cost of Developing Large-Scale Software, *IEEE Transactions on Computers*, Vol. 23, No. 6, pp. 615-636.

Wasserman, P. D., 1989, *Neural Computing: Theory and Practice*, Van Norstand Reinhold, New York.

Weimer, J., 1995, *Research Techniques in Human Engineering*, Prentice Hall, Englewood Cliffs, New Jersey.

# Appendix I: Project Data

The data collected from the 32 historical projects from NN, CMC, and GE are summarized in this appendix.

### I.1 Project Data for NN

| Project number | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| **Actual design effort (hours)** | 4616 | 8800 | 7500 | 11468 | 25033 |
| | | | | | |
| **Estimated design effort (hours)** | 1300 | 4700 | 6850 | 13500 | 29500 |
| | | | | | |
| **Estimation Method** | | | | | |
| • Expert judgement | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Metrics approach | | | | | |
| • Other | | | | | |
| | | | | | |
| **The estimation was done by** | | | | | |
| • Project managers | | | | | |
| • Design managers | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Designers | | | | | |
| • Other | | | | | |
| | | | | | |
| **Project schedule tool** | | | | | |
| • CPM | | | | | |
| • PERT | | | | | |
| • GANTT chart | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Other | | | | | |

## I.1 Project Data for NN (continued)

| Project number | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| **Premium on early completion** | | | | | |
| • Low | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Medium | | | | | |
| • High | | | | | |
| | | | | | |
| **Type of process used** | | | | | |
| • No process used | | | | | |
| • Phase review process | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Other | | | | | |
| | | | | | |
| **Use of computer assisted tools** | | | | | |
| • No tools have been used | | | | | |
| • Tools have been used for drawings only | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Tools have been used in most of the design phases | | | | | |
| | | | | | |
| **Project progress measurement** | | | | | |
| • Subjective statements | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Metrics | | | | | |
| • Other | | | | | |
| • | | | | | |
| **Project monitoring method** | | | | | |
| • Written report | | | | | |
| • Formal meetings | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Informal meetings | | | | | |
| • Other | | | | | |

## I.1 Project Data for NN (continued)

| Project number | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| **Monitoring scheduling policies** | | | | | |
| • At regular intervals | ✓ | ✓ | ✓ | ✓ | ✓ |
| • At random intervals | | | | | |
| • Other | | | | | |
| | | | | | |
| **Method of communication for design team** | | | | | |
| • Decentralized | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Centralized | | | | | |
| | | | | | |
| **Type of designed product** | | | | | |
| • Mechanical | | | | | |
| • Electronic / Electrical | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Software | | | | | |
| • Other | | | | | |
| | | | | | |
| **Product complexity** | 43 | 73 | 76 | 90 | 135 |
| | | | | | |
| **Severity of requirements** | | | | | |
| 1: design requirements were not too difficult to meet | | | | | |
| 2: design requirements were difficult to meet | | ✓ | ✓ | | |
| 3: design requirements were extremely difficult to meet | ✓ | | | ✓ | ✓ |

## I.2 Project Data for CMC

| Project number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Actual design effort (hours)** | 951 | 632 | 1103 | 1099 | 1367 | 1874 | 1985 | 1777 | 4950 | 3701 | 8883 | 9828 |
| **Estimated design effort (hours)** | 983 | 478 | 364 | 946 | 1223 | 3232 | 566 | 831 | 3716 | 1106 | 12480 | 13898 |
| **Estimation Method** | | | | | | | | | | | | |
| • Expert judgement | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Metrics approach | | | | | | | | | | | | |
| • Other | | | | | | | | | | | | |
| **The estimation was done by** | | | | | | | | | | | | |
| • Project managers | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Design managers | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Designers | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Other | | | | | | | | | | | | |
| **Project schedule tool** | | | | | | | | | | | | |
| • CPM | | | | | | | | | | | | |
| • PERT | | | | | | | | | | | | |
| • GANTT chart | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Other | | | | | | | | | | | | |
| **Premium on early completion** | | | | | | | | | | | | |
| • Low | | | | | | | | | | | | |
| • Medium | | | | | | | | | | | | |
| • High | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| **Type of process used** | | | | | | | | | | | | |
| • No process used | | | | | | | | | | | | |
| • Phase review process | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Other | | | | | | | | | | | | |

## I.2 Project Data for CMC (continued)

| Project number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Use of computer assisted tools** | | | | | | | | | | | | |
| • No tools have been used | | | | | | | | | | | | |
| • Tools have been used for drawings only | | | | | | | | | | | | |
| • Tools have been used in most of the design phases | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| | | | | | | | | | | | | |
| **Project progress measurement** | | | | | | | | | | | | |
| • Subjective statements | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Metrics | | | | | | | | | | | | |
| • Other | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| **Project monitoring method** | | | | | | | | | | | | |
| • Written report | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Formal meetings | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Informal meetings | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Other | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| **Monitoring scheduling policies** | | | | | | | | | | | | |
| • At regular intervals | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • At random intervals | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Other | | | | | | | | | | | | |

## I.2 Project Data for CMC (Continued)

| Project number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Method of communication for design team** | | | | | | | | | | | | |
| • Decentralized | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Centralized | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| **Type of designed product** | | | | | | | | | | | | |
| • Mechanical | | | | | | | | | | | | |
| • Electronic/ Electrical | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Software | | | | | | | | | | | | |
| • Other | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| **Product complexity** | 5 | 7 | 7 | 11 | 11 | 11 | 11 | 15 | 19 | 22 | 24 | 34 |
| | | | | | | | | | | | | |
| **Severity of requirements** | | | | | | | | | | | | |
| 1: design requirements were not too difficult to meet | | ✓ | | ✓ | | | | ✓ | | | | |
| 2: design requirements were difficult to meet | ✓ | | ✓ | | ✓ | ✓ | ✓ | | ✓ | ✓ | | |
| 3: design requirements were extremely difficult to meet | | | | | | | | | | | ✓ | ✓ |

## I.3 Project Data for GE

| Project number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| **Actual design effort (hours)** | 20392 | 8192 | 13544 | 11880 | 8384 | 27200 | 20800 | 30400 |
| | | | | | | | | |
| **Estimated design effort (hours)** | 19600 | 11600 | 13088 | 14000 | 11200 | 15400 | 16600 | 20000 |
| | | | | | | | | |
| **Actual engineering effort (hours)** | 8762 | 2160 | 2544 | 2080 | 1680 | 6904 | 4616 | 10048 |
| | | | | | | | | |
| **Estimated engineering effort (hours)** | 6800 | 2000 | 3200 | 2800 | 1600 | 7200 | 3000 | 4000 |
| | | | | | | | | |
| **Estimation Method** | | | | | | | | |
| • Expert judgement | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Metrics approach | | | | | | | | |
| • Other | | | | | | | | |
| | | | | | | | | |
| • **The estimation was done by** | | | | | | | | |
| • Project managers | | | | | | | | |
| • Design managers | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Designers | | | | | | | | |
| • Other | | | | | | | | |
| | | | | | | | | |
| **Project schedule tool** | | | | | | | | |
| • CPM | | | | | | | | |
| • PERT | | | | | | | | |
| • GANTT chart | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Other | | | | | | | | |
| | | | | | | | | |
| **Premium on early completion** | | | | | | | | |
| • Low | | | | | | | | |
| • Medium | | | | | | | | |
| • High | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

## I.3 Project Data for GE (Continued)

| Project number | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|
| **Actual design effort (hours)** | 19824 | 16944 | 20112 | 26814 | 10704 | 10856 | 8760 |
| | | | | | | | |
| **Estimated design effort (hours)** | 17400 | 17400 | 16000 | 24000 | 13600 | 16000 | 13200 |
| | | | | | | | |
| **Actual engineering effort (hours)** | 5592 | 5272 | 5776 | 10008 | 1520 | 2816 | 1288 |
| | | | | | | | |
| **Estimated engineering effort (hours)** | 3600 | 3600 | 4000 | 8000 | 2800 | 3200 | 2000 |
| | | | | | | | |
| **Estimation Method** | | | | | | | |
| • Expert judgement | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Metrics approach | | | | | | | |
| • Other | | | | | | | |
| | | | | | | | |
| **The estimation was done by** | | | | | | | |
| • Project managers | | | | | | | |
| • Design managers | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Designers | | | | | | | |
| • Other | | | | | | | |
| | | | | | | | |
| **Project schedule tool** | | | | | | | |
| • CPM | | | | | | | |
| • PERT | | | | | | | |
| • GANTT chart | | | | | | | |
| • Other | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| | | | | | | | |
| **Premium on early completion** | | | | | | | |
| • Low | | | | | | | |
| • Medium | | | | | | | |
| • High | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

## I.3 Project Data for GE (continued)

| Project number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| **Type of process used** | | | | | | | | |
| • No process used | | | | | | | | |
| • Phase review process | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Other | | | | | | | | |
| | | | | | | | | |
| **Use of computer assisted tools** | | | | | | | | |
| • No tools have been used | | | | | | | | |
| • Tools have been used for drawings only | | | | | | | | |
| • Tools have been used in most of the design phases | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| | | | | | | | | |
| **Project progress measurement** | | | | | | | | |
| • Subjective statements | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Metrics | | | | | | | | |
| • Other | | | | | | | | |
| | | | | | | | | |
| **Project monitoring method** | | | | | | | | |
| • Written report | | | | | | | | |
| • Formal meetings | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Informal meetings | | | | | | | | |
| • Other | | | | | | | | |
| | | | | | | | | |
| **Monitoring scheduling policies** | | | | | | | | |
| • At regular intervals | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • At random intervals | ✓ | | | | | | | |
| • Other | | | | | | | | |

## I.3 Project Data for GE (continued)

| Project number | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|
| **Type of process used** | | | | | | | |
| • No process used | | | | | | | |
| • Phase review process | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Other | | | | | | | |
| | | | | | | | |
| **Use of computer assisted tools** | | | | | | | |
| • No tools have been used | | | | | | | |
| • Tools have been used for drawings only | | | | | | | |
| • Tools have been used in most of the design phases | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| | | | | | | | |
| **Project progress measurement** | | | | | | | |
| • Subjective statements | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Metrics | | | | | | | |
| • Other | | | | | | | |
| | | | | | | | |
| **Project monitoring method** | | | | | | | |
| • Written report | | | | | | | |
| • Formal meetings | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Informal meetings | | | | | | | |
| • Other | | | | | | | |
| | | | | | | | |
| **Monitoring scheduling policies** | | | | | | | |
| • At regular intervals | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • At random intervals | | | | | | | |
| • Other | | | | | | | |

## I.3 Project Data for GE (Continued)

| Project number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| **Method of communication for design team** | | | | | | | | |
| • Decentralized | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Centralized | | | | | | | | |
| | | | | | | | | |
| **Type of designed product** | | | | | | | | |
| • Mechanical | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Electronic/ Electrical | | | | | | | | |
| • Software | | | | | | | | |
| • Other | | | | | | | | |
| | | | | | | | | |
| **Product complexity** | 383 | 322 | 322 | 308 | 3319 | 368 | 322 | 352 |
| | | | | | | | | |
| **Technical difficulty to expertise ratio** | 1.2 | 0.6 | 0.9 | 0.9 | 0.5 | 1.0 | 0.8 | 1.1 |
| | | | | | | | | |
| **Type of drawings submitted to the customer** | | | | | | | | |
| 1: basic drawings | ✓ | ✓ | ✓ | ✓ | ✓ | | | |
| 2: assembly drawings | | | | | | ✓ | ✓ | ✓ |
| 3: manufacturing level drawings | | | | | | | | |
| | | | | | | | | |
| **Involvement of design partners** | | | | | | | | |
| 1: no design partners were involved | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | |
| 2: design partners were involved | | | | | | | | ✓ |

## I.3 Project Data for GE (Continued)

| **Project number** | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|
| **Method of communication for design team** | | | | | | | |
| • Decentralized | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Centralized | | | | | | | |
| | | | | | | | |
| **Type of designed product** | | | | | | | |
| • Mechanical | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Electronic/ Electrical | | | | | | | |
| • Software | | | | | | | |
| • Other | | | | | | | |
| | | | | | | | |
| **Product complexity** | 383 | 322 | 322 | 308 | 3319 | 368 | 322 |
| | | | | | | | |
| **Technical difficulty to expertise ratio** | 1.0 | 1.1 | 1.0 | 1.1 | 0.7 | 0.9 | 0.8 |
| | | | | | | | |
| **Type of drawings submitted to the customer** | | | | | | | |
| 1: basic drawings | | | | | ✓ | ✓ | ✓ |
| 2: assembly drawings | ✓ | ✓ | | ✓ | | | |
| 3: manufacturing level drawings | | | ✓ | | | | |
| | | | | | | | |
| **Involvement of design partners** | | | | | | | |
| 1: no design partners were involved | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ |
| 2: design partners were involved | | | | ✓ | | | |

# Appendix II: Design Tasks

This appendix contains the instructions given to the subjects who participated in the experiment and the three design tasks, A, B, and C as discussed in Chapter 4.

## II.1 Instructions

Dear Participant

- Before starting the experiment, you will be given an unlimited time training session. This will help in understanding how to use the provided tables and documentation. Training time will be excluded from the experiment duration.
- You are allowed to take breaks whenever you begin to feel tired. Break time is excluded from the experiment time.
- The use of materials and standard components other than those in the provided references is not allowed.
- There is no time limit to complete your task.
- Work independently.
- Time, simplicity, and meeting the requirements are very important factors. Try to make your design simple and meet all the requirements in minimum time.
- Use the provided checklist to be sure that you have met all the requirements.

Thank you for your cooperation

## II.2 Task A

A device to be designed that is capable of positioning a workpiece at a height of 30 in. (76.2 cm).

List of requirements

- Working height: 30 in. (76.2 cm) above floor.
- For clamping purpose, the surface must have at least 6 equally spaced holes of 1.0± 0.0005 in. (2.54 ± 0.001 cm) diameter.
- Accessibility: weld locations over whole surface (good accessibility from all sides).
- Workpiece
  - Material: steel, steel castings
  - Size: maximum base 20 × 20 in. (50.8 × 50.8 cm)
  - Mass: maximum 50 lbs (22.68 kg)
- Safety
  - The device should be firmly fixed to the floor so there is no chance for it turning over or slipping accidentally.
- Maintenance
  - Maintenance requirement: minimum
- Manufacture
  - Small batch

The assignment

1. Given the above requirements, you are required to design a product whose functional tree is shown in Figure II.1.
2. Produce an assembly drawing (the front and top views) of the product including leading dimensions.
   - This drawing should adequately define the geometry of all parts.
   - Drawings should be made to the scale.
   - The individual components should be numbered on the drawing and listed in a bill of materials. The bill of materials should include the materials used. Also included should be any standard components along with their reference number.

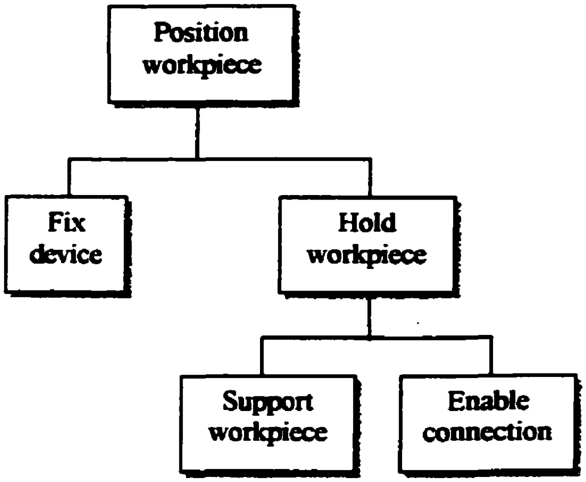



**Figure II.1** Functional decomposition for task A

## II.3 Task B

A device to be designed that is capable of positioning a workpiece at a desired height.

List of requirements

- All the necessary movements to be manual operations.
- The desired position can be obtained by adjusting height between 30 in. to 50 in. (76.2 to 127 cm)
- For clamping purpose, the surface must have at least 6 equally spaced holes of 1.0± 0.0005 in. (2.54 ± 0.001 cm) diameter.
- Accessibility: weld locations over whole surface (good accessibility from all sides).
- Workpiece
    - Material: steel, steel castings
    - Size: maximum base 20 × 20 in. (50.8 × 50.8 cm)
    - Mass: maximum 50 lbs (22.68 kg)
- Safety

    - The device should be firmly fixed to the floor so there is no chance for it turning over or slipping accidentally.

- Maintenance
    - Maintenance requirement: minimum
- Manufacture
    - Small batch

The assignment

1. Given the above requirements, you are required to design a product whose functional tree is shown in Figure II.2.
2. Produce an assembly drawing (the front and top views) of the product including leading dimensions.
    - This drawing should adequately define the geometry of all parts.
    - Drawings should be made to scale.
    - The individual components should be numbered on the drawing and listed in a bill of materials. The bill of materials should include the materials used. Also included should be any standard components along with their reference number.

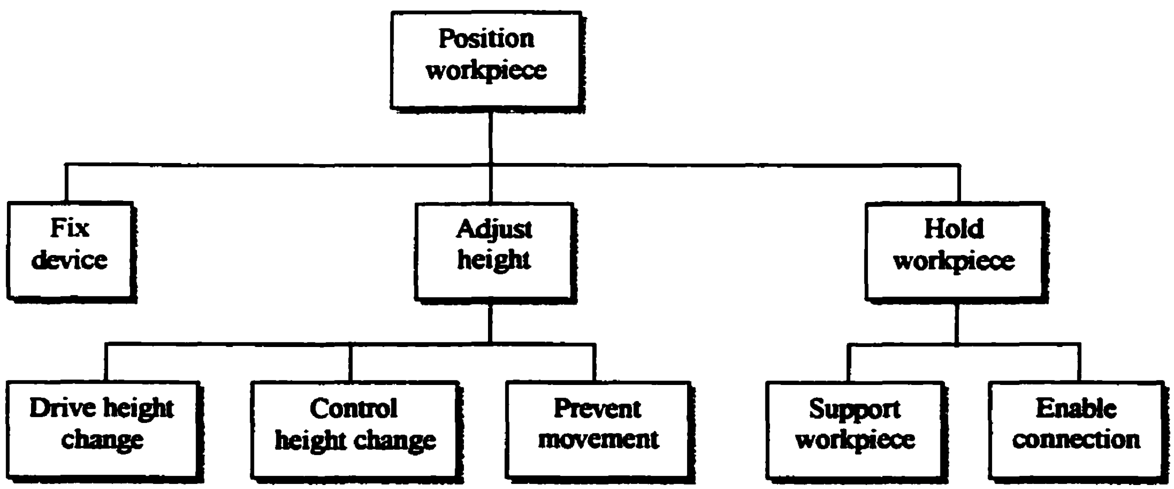



**Figure II.2** Functional decomposition for task B

## II.4 Task C

A welding positioner is to be designed that is capable of positioning a workpiece at a desired position for welding.

List of requirements

- All the necessary movements to be manual operations.
- The desired position can be obtained by adjusting height between 30 in. to 50 in. (76.2 to 127 cm), and rotating about vertical axis.
- For clamping purpose, the surface must have at least 6 equally spaced holes of 1.0± 0.0005 in. (2.54 ± 0.001 cm) diameter.
- Accessibility: weld locations over whole surface (good accessibility from all sides).
- Workpiece
  - Material: steel, steel castings
  - Size: maximum base 20 × 20 in. (50.8 × 50.8 cm)
  - Mass: maximum 50 lbs (22.68 kg)
- Safety
  - The device should be firmly fixed to the floor so there is no chance for it turning over or slipping accidentally.
- Maintenance
  - Maintenance requirement: minimum
- Manufacture
  - Small batch

The assignment

1. Given the above requirements, you are required to design a product whose functional tree is shown in Figure II.3.
2. Produce an assembly drawing (the front and top views) of the product including leading dimensions.
   - This drawing should adequately define the geometry of all parts.
   - Drawings should be made to scale.

- The individual components should be numbered on the drawing and listed in a bill of materials. The bill of materials should include the materials used. Also included should be any standard components along with their reference number.

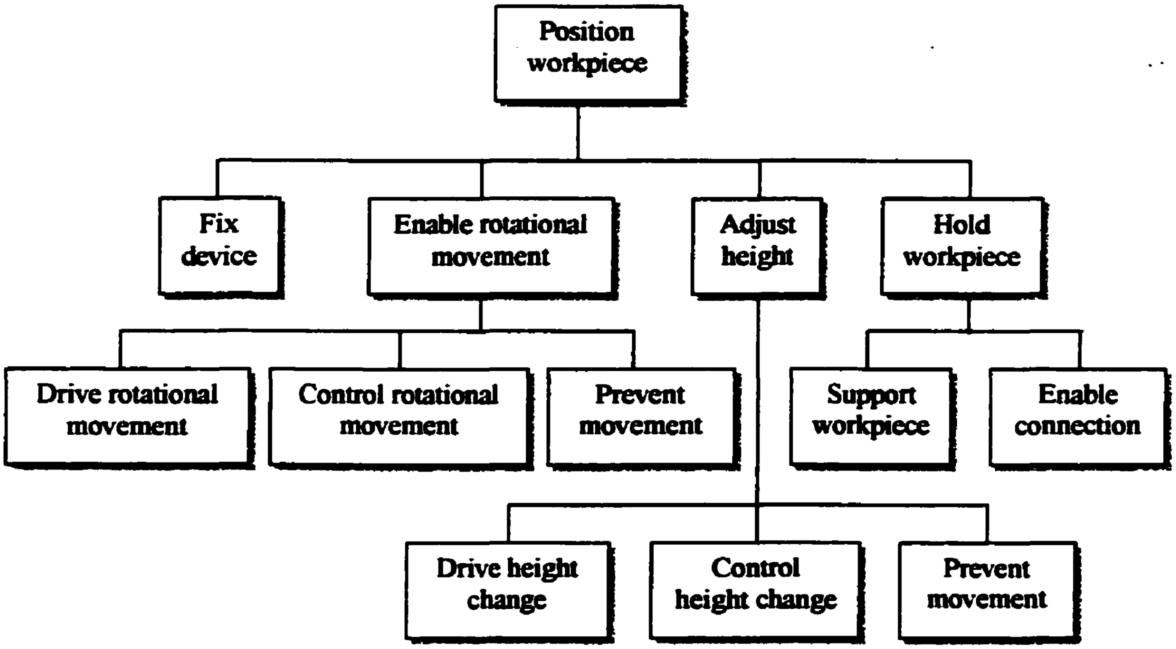



**Figure II.3** Functional decomposition for task C

# Appendix III: Use of the Principal Right Eigenvector in the Analogy-Based Model

In this appendix, the basic mathematical reasoning underlying the use of the principal right eigenvector in the analogy-based model is explained. For a more thorough treatment of this issue, the reader is referred to Saaty (1980).

The objective of using the eigenvector approach in the analogy-based model is to estimate the weights of influence of a factor, $f$, on the productivity of a set of projects (an upcoming project and a number of reference projects) from a matrix of pairwise comparisons, $A = (a_{ij})$. These weights are then used to compute the multipliers, $M_{rf}$. Thus, given the matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1s} \\ a_{21} & a_{22} & \cdots & a_{2s} \\ \vdots & \vdots & \vdots & \vdots \\ a_{s1} & a_{s2} & \cdots & a_{ss} \end{bmatrix}$$

where $a_{ij} = \dfrac{1}{a_{ji}}$ for all $i,\ j = 1, 2, \ldots, s$, a vector of weights, $w = (w_1, w_2, \ldots, w_s)$, needs to be computed. If the judgements were perfectly consistent, i.e., $a_{ij} = a_{ik}a_{kj}$ for all $i, j, k = 1, 2, \ldots, s$, then, the entries of matrix $A$ would contain no errors and could be expressed as:

$$a_{ij} = \frac{w_i}{w_j} \qquad i, j = 1, 2, \ldots, s$$

and thus

$$a_{ij}\frac{w_j}{w_i}=1 \qquad i,j=1,2,\ldots,s$$

Consequently

$$\sum_{j=1}^{s}a_{ij}w_j\frac{1}{w_i}=s \qquad i=1,2,\ldots,s$$

or

$$\sum_{j=1}^{s}a_{ij}w_j=s\,w_i \qquad i=1,2,\ldots,s$$

which is equivalent to

$$A\,w=s\,w \tag{III.1}$$

In matrix theory, if $\lambda_1,\cdots,\lambda_s$ are the numbers satisfying equation (III.2), i.e., are the eigenvalues of $A$, and if $a_{ii}=1$ for all $i$, then $\sum_{i=1}^{s}\lambda_i=s$ (Heesterman, 1990).

$$A\,x=\lambda\,x \tag{III.2}$$

Therefore, if equation (III.1) holds, then all eigenvalues are zero, except one, which is $s$. Clearly, then, in the consistent case, $s$ is the largest eigenvalue of $A$, and $w$ is the principal right eigenvector. Furthermore, if the entries $a_{ij}$ are changed by small amounts, then the eigenvalues change by small amounts. In other words, if the diagonal of a matrix $A$ consists of ones, $a_{ii}=1$, and if $A$ is consistent, then small variations of the $a_{ij}$ elements keep the largest eigenvalue, $\lambda_{\max}$, close to $s$, and the remaining eigenvalues close to zero. Thus, the deviation of $\lambda_{\max}$ from $s$ can provide a measure of consistency. On this basis, Saaty (1980) proposed the consistency index, *CI*, presented in Chapter 7 and given in equation (III.3).

$$CI = \frac{\lambda_{\max} - s}{s - 1} \tag{III.3}$$

# Appendix IV: Analogy-Based Model Computations

## IV.1 Analogy-Based Model Computations for NN

| Upcoming project | Reference project | $P_r$ | $M_{rf}$ | $P_{ur}$ | $E_{ur}$ | $E_u$ |
|---|---|---|---|---|---|---|
| 1 | 2 | 1.26 | 0.92 | 1.16 | 37.07 | 31.39 |
| | 3 | 1.54 | 0.83 | 1.28 | 33.59 | |
| | 4 | 1.19 | 1.42 | 1.69 | 25.44 | |
| | 5 | 0.82 | 1.78 | 1.46 | 29.45 | |
| 2 | 1 | 1.42 | 1.09 | 1.55 | 47.10 | 46.21 |
| | 3 | 1.54 | 0.91 | 1.40 | 52.14 | |
| | 4 | 1.19 | 1.55 | 1.84 | 39.67 | |
| | 5 | 0.82 | 1.94 | 1.59 | 45.91 | |
| 3 | 1 | 1.42 | 1.20 | 1.70 | 44.71 | 45.11 |
| | 2 | 1.26 | 1.10 | 1.39 | 54.68 | |
| | 4 | 1.19 | 1.70 | 2.02 | 37.62 | |
| | 5 | 0.82 | 2.13 | 1.75 | 43.43 | |
| 4 | 1 | 1.42 | 0.71 | 1.01 | 89.11 | 96.29 |
| | 2 | 1.26 | 0.65 | 0.82 | 109.76 | |
| | 3 | 1.54 | 0.59 | 0.91 | 98.90 | |
| | 5 | 0.82 | 1.25 | 1.03 | 87.38 | |
| 5 | 1 | 1.42 | 0.56 | 0.80 | 168.75 | 175.98 |
| | 2 | 1.26 | 0.52 | 0.66 | 206.55 | |
| | 3 | 1.54 | 0.47 | 0.72 | 187.50 | |
| | 4 | 1.19 | 0.80 | 0.95 | 141.11 | |

## IV.2 Analogy-Based Model Computations for CMC

| Upcoming project | Reference project | $P_r$ | $M_{rf}$ | $P_{uri}$ | $E_{ur}$ | $E_u$ |
|---|---|---|---|---|---|---|
| 1 | 2 | 1.68 | 0.54 | 0.91 | 5.51 | 4.81 |
| | 3 | 0.96 | 0.97 | 0.93 | 5.37 | |
| | 4 | 1.52 | 0.65 | 0.99 | 5.06 | |
| | 5 | 1.22 | 0.90 | 1.10 | 4.55 | |
| | 6 | 0.89 | 1.05 | 0.93 | 5.35 | |
| | 7 | 0.84 | 1.41 | 1.18 | 4.22 | |
| | 8 | 1.28 | 0.68 | 0.87 | 5.74 | |
| | 9 | 0.58 | 1.71 | 0.99 | 5.04 | |
| | 10 | 0.9 | 1.40 | 1.26 | 3.97 | |
| | 11 | 0.41 | 2.88 | 1.18 | 4.23 | |
| | 12 | 0.53 | 2.43 | 1.29 | 3.88 | |
| 2 | 1 | 0.8 | 1.84 | 1.47 | 4.76 | 3.71 |
| | 3 | 0.96 | 1.78 | 1.71 | 4.10 | |
| | 4 | 1.52 | 1.20 | 1.82 | 3.84 | |
| | 5 | 1.22 | 1.65 | 2.01 | 3.48 | |
| | 6 | 0.89 | 1.94 | 1.73 | 4.05 | |
| | 7 | 0.84 | 2.59 | 2.18 | 3.22 | |
| | 8 | 1.28 | 1.26 | 1.61 | 4.34 | |
| | 9 | 0.58 | 3.15 | 1.83 | 3.83 | |
| | 10 | 0.9 | 2.58 | 2.32 | 3.01 | |
| | 11 | 0.41 | 5.30 | 2.17 | 3.22 | |
| | 12 | 0.53 | 4.47 | 2.37 | 2.95 | |
| 3 | 1 | 0.8 | 1.03 | 0.82 | 8.50 | 6.63 |
| | 2 | 1.68 | 0.56 | 0.94 | 7.44 | |
| | 4 | 1.52 | 0.67 | 1.02 | 6.87 | |
| | 5 | 1.22 | 0.92 | 1.12 | 6.24 | |
| | 6 | 0.89 | 1.09 | 0.97 | 7.22 | |
| | 7 | 0.84 | 1.45 | 1.22 | 5.75 | |
| | 8 | 1.28 | 0.71 | 0.91 | 7.70 | |
| | 9 | 0.58 | 1.77 | 1.03 | 6.82 | |
| | 10 | 0.9 | 1.45 | 1.31 | 5.36 | |
| | 11 | 0.41 | 2.97 | 1.22 | 5.75 | |
| | 12 | 0.53 | 2.51 | 1.33 | 5.26 | |
| 4 | 1 | 0.8 | 1.54 | 1.23 | 8.93 | 7.04 |
| | 2 | 1.68 | 0.83 | 1.39 | 7.89 | |
| | 3 | 0.96 | 1.49 | 1.43 | 7.69 | |
| | 5 | 1.22 | 1.38 | 1.68 | 6.53 | |
| | 6 | 0.89 | 1.62 | 1.44 | 7.63 | |
| | 7 | 0.84 | 2.16 | 1.81 | 6.06 | |
| | 8 | 1.28 | 1.05 | 1.34 | 8.18 | |
| | 9 | 0.58 | 2.63 | 1.53 | 7.21 | |
| | 10 | 0.9 | 2.15 | 1.94 | 5.68 | |
| | 11 | 0.41 | 4.42 | 1.81 | 6.07 | |
| | 12 | 0.53 | 3.73 | 1.98 | 5.56 | |

## IV.2 Analogy-Based Model Computations for CMC (continued)

| Upcoming project | Reference project | $P_r$ | $M_{rf}$ | $P_{ur}$ | $E_{ur}$ | $E_u$ |
|---|---|---|---|---|---|---|
| 5 | 1 | 0.8 | 1.12 | 0.90 | 12.28 | 9.77 |
| | 2 | 1.68 | 0.61 | 1.02 | 10.73 | |
| | 3 | 0.96 | 1.08 | 1.04 | 10.61 | |
| | 4 | 1.52 | 0.73 | 1.11 | 9.91 | |
| | 6 | 0.89 | 1.18 | 1.05 | 10.47 | |
| | 7 | 0.84 | 1.57 | 1.32 | 8.34 | |
| | 8 | 1.28 | 0.76 | 0.97 | 11.31 | |
| | 9 | 0.58 | 1.91 | 1.11 | 9.93 | |
| | 10 | 0.9 | 1.56 | 1.40 | 7.83 | |
| | 11 | 0.41 | 3.22 | 1.32 | 8.33 | |
| | 12 | 0.53 | 2.71 | 1.44 | 7.66 | |
| 6 | 1 | 0.8 | 0.95 | 0.76 | 14.47 | 11.33 |
| | 2 | 1.68 | 0.52 | 0.87 | 12.59 | |
| | 3 | 0.96 | 0.92 | 0.88 | 12.45 | |
| | 4 | 1.52 | 0.62 | 0.94 | 11.67 | |
| | 5 | 1.22 | 0.85 | 1.04 | 10.61 | |
| | 7 | 0.84 | 1.34 | 1.13 | 9.77 | |
| | 8 | 1.28 | 0.65 | 0.83 | 13.22 | |
| | 9 | 0.58 | 1.62 | 0.94 | 11.71 | |
| | 10 | 0.9 | 1.33 | 1.20 | 9.19 | |
| | 11 | 0.41 | 2.74 | 1.12 | 9.79 | |
| | 12 | 0.53 | 2.31 | 1.22 | 8.98 | |
| 7 | 1 | 0.8 | 0.71 | 0.57 | 19.37 | 15.47 |
| | 2 | 1.68 | 0.39 | 0.66 | 16.79 | |
| | 3 | 0.96 | 0.69 | 0.66 | 16.61 | |
| | 4 | 1.52 | 0.46 | 0.70 | 15.73 | |
| | 5 | 1.22 | 0.64 | 0.78 | 14.09 | |
| | 6 | 0.89 | 0.75 | 0.67 | 16.48 | |
| | 8 | 1.28 | 0.49 | 0.63 | 17.54 | |
| | 9 | 0.58 | 1.21 | 0.70 | 15.67 | |
| | 10 | 0.9 | 0.99 | 0.89 | 12.35 | |
| | 11 | 0.41 | 2.04 | 0.84 | 13.15 | |
| | 12 | 0.53 | 1.72 | 0.91 | 12.07 | |
| 8 | 1 | 0.8 | 1.46 | 1.17 | 12.84 | 9.97 |
| | 2 | 1.68 | 0.79 | 1.33 | 11.30 | |
| | 3 | 0.96 | 1.42 | 1.36 | 11.00 | |
| | 4 | 1.52 | 0.95 | 1.44 | 10.39 | |
| | 5 | 1.22 | 1.31 | 1.60 | 9.39 | |
| | 6 | 0.89 | 1.54 | 1.37 | 10.94 | |
| | 7 | 0.84 | 2.06 | 1.73 | 8.67 | |
| | 9 | 0.58 | 2.50 | 1.45 | 10.34 | |
| | 10 | 0.9 | 2.05 | 1.85 | 8.13 | |
| | 11 | 0.41 | 4.21 | 1.73 | 8.69 | |
| | 12 | 0.53 | 3.55 | 1.88 | 7.97 | |

## IV.2 Analogy-Based Model Computations for CMC (continued)

| Upcoming project | Reference project | $P_r$ | $M_{rf}$ | $P_{ur}$ | $E_{ur}$ | $E_u$ |
|---|---|---|---|---|---|---|
| 9 | 1 | 0.8 | 0.58 | 0.46 | 40.95 | 31.96 |
| | 2 | 1.68 | 0.32 | 0.54 | 35.34 | |
| | 3 | 0.96 | 0.57 | 0.55 | 34.72 | |
| | 4 | 1.52 | 0.38 | 0.58 | 32.89 | |
| | 5 | 1.22 | 0.52 | 0.63 | 29.95 | |
| | 6 | 0.89 | 0.62 | 0.55 | 34.43 | |
| | 7 | 0.84 | 0.82 | 0.69 | 27.58 | |
| | 8 | 1.28 | 0.40 | 0.51 | 37.11 | |
| | 10 | 0.9 | 0.82 | 0.74 | 25.75 | |
| | 11 | 0.41 | 1.68 | 0.69 | 27.58 | |
| | 12 | 0.53 | 1.42 | 0.75 | 25.25 | |
| 10 | 1 | 0.8 | 0.71 | 0.57 | 38.73 | 30.90 |
| | 2 | 1.68 | 0.39 | 0.66 | 33.58 | |
| | 3 | 0.96 | 0.69 | 0.66 | 33.21 | |
| | 4 | 1.52 | 0.46 | 0.70 | 31.46 | |
| | 5 | 1.22 | 0.64 | 0.78 | 28.18 | |
| | 6 | 0.89 | 0.75 | 0.67 | 32.96 | |
| | 7 | 0.84 | 1.01 | 0.85 | 25.93 | |
| | 8 | 1.28 | 0.49 | 0.63 | 35.08 | |
| | 9 | 0.58 | 1.22 | 0.71 | 31.09 | |
| | 11 | 0.41 | 2.06 | 0.84 | 26.05 | |
| | 12 | 0.53 | 1.73 | 0.92 | 23.99 | |
| 11 | 1 | 0.8 | 0.35 | 0.28 | 85.71 | 68.99 |
| | 2 | 1.68 | 0.19 | 0.32 | 75.19 | |
| | 3 | 0.96 | 0.34 | 0.33 | 73.53 | |
| | 4 | 1.52 | 0.23 | 0.35 | 68.65 | |
| | 5 | 1.22 | 0.31 | 0.38 | 63.46 | |
| | 6 | 0.89 | 0.37 | 0.33 | 72.88 | |
| | 7 | 0.84 | 0.49 | 0.41 | 58.31 | |
| | 8 | 1.28 | 0.24 | 0.31 | 78.13 | |
| | 9 | 0.58 | 0.59 | 0.34 | 70.13 | |
| | 10 | 0.9 | 0.49 | 0.44 | 54.42 | |
| | 12 | 0.53 | 0.84 | 0.45 | 53.91 | |
| 12 | 1 | 0.8 | 0.41 | 0.33 | 103.66 | 82.95 |
| | 2 | 1.68 | 0.22 | 0.37 | 91.99 | |
| | 3 | 0.96 | 0.40 | 0.38 | 88.54 | |
| | 4 | 1.52 | 0.27 | 0.41 | 82.85 | |
| | 5 | 1.22 | 0.37 | 0.45 | 75.32 | |
| | 6 | 0.89 | 0.43 | 0.38 | 88.84 | |
| | 7 | 0.84 | 0.58 | 0.49 | 69.79 | |
| | 8 | 1.28 | 0.28 | 0.36 | 94.87 | |
| | 9 | 0.58 | 0.70 | 0.41 | 83.74 | |
| | 10 | 0.9 | 0.58 | 0.52 | 65.13 | |
| | 11 | 0.41 | 1.19 | 0.49 | 69.69 | |