43. ANALOGY BASED ESTIMATION IN BUILDING SERVICES

Kai Rintala[1], Bimal Kumar[2], Iain MacLeod[2]

[1]Bartlett School of Graduate Studies, University College London
[2]Department of Civil Engineering, University of Strathclyde

Abstract

The introduction of Private Finance Initiative in the United Kingdom has prompted the need for quality whole life cost information. This in turn has led to a search for new innovative methods for producing whole life cost information. The research reported in this paper is a part of an overall approach being developed for whole life costing of building services. A particular problem identified in producing initial estimates of the whole life cost of building services is that the information on the building services design is limited. Analogy Based Estimation (ABE) was investigated as a potential approach to generate more information for estimation purposes and to estimate the capital cost of building services. The performance of ABE was investigated in three test settings. The estimation performance of ABE was compared with the estimation performance of linear regression (LR). ABE outperformed LR on two test settings. However, the estimation performance of ABE was found not to be sufficient for the approach to be applied in industry prior to further investigation. The tests were hindered by lack of data and inconsistencies in the data obtained. Therefore, no conclusions can be made on the applicability of ABE in this particular problem domain and ABE remains a potential approach. The research described in this paper will stand as a starting point for further testing and development of ABE in the estimation of building services.

Keywords: *analogy based estimation, building services, capital cost*

RATIONALE FOR APPROACH

The introduction of Private Finance Initiative in the United Kingdom has prompted the need for quality whole life cost (WLC) information. This in turn has led to a search for new innovative methods for producing whole life cost information. One particular problem identified in whole life costing of building services is that initial estimates need to be produced based on incomplete information on the design. The document reports on an investigation of a potential solution to this problem.

The search for a potential solution focused on artificial intelligence (AI) techniques. This was because some AI based solutions are known to be most beneficial in problem domains where the information available is limited or incomplete (Ritter, 1980). The initial WLC estimates of building services are needed in such a context.

Analogy based estimation (ABE) was chosen to be investigated as a potential solution due to three specific characteristics of the technique. Firstly, ABE has been found to have a good configurability, which is the effort needed to produce useful results with a technique (Mair et al, 2000). Secondly, the use of a technique that resembles the actual reasoning process in the problem domain has been found beneficial (Marir and Watson, 1995). The ABE process closely resembles the reasoning process that estimators use in the estimation of the capital cost of building services. This makes the ABE process easy for estimators to relate to. Thirdly, the ABE process is readily understandable as distinct form e.g. artificial neural networks.

In addition to the favourable characteristics, Schofield and Shepperd (1995) have highlighted two weaknesses in the approach. These are the inability to cope with very small sets of data (i.e. insufficient number analogies) and the inability to cope in circumstances of great variation (i.e. insufficient number of appropriate analogies).

ABE is widely used within the software industry. It has been applied primarily in estimation of the effort required in software production projects (Cowderoy and Jenkins, 1988).

WHAT IS ANALOGY BASED ESTIMATION?

The existing body of literature does not define the difference between ABE and estimation by case based reasoning (CBR). Riesbeck and Schank (1989) define CBR to be *'a process of solving new problems by adapting solutions that were used to solve old problems.'* Aarmodt and Plaza (1994) describe CBR as a cyclic process composed of retrieval of similar cases, reuse of the retrieved cases to find a solution to the problem, adaptation of the proposed solution if necessary, and retention of the solution to form a new case. According to Shepperd *et al* (1996), the ABE process includes the characterisation of the project for which an estimate is required, finding similar cases for which the target value is known, and using values from the similar cases to generate an estimate.

The difference in the descriptions of CBR and ABE processes given in the previous paragraph is the retention of solutions for future use. Shepperd *et al* (1996) have developed a tool called ANGEL for estimation of software production effort using ABE. Despite the differences in the descriptions of CBR and ABE processes, the ANGEL tool has a facility to retain generated solutions for future use. The retention of solved estimation problems gives ABE the capability to learn from past experience.

Mair *et al* (2000) have used the terms *ABE* and *CBR* interchangeably. Therefore, it could be concluded that ABE is estimation by CBR. However, in the literature there is another apparent pattern in the use of these terms. Schofield and Shepperd (1995), who use the term *ABE*, use algorithms to identify the closest analogies and to generate estimates. The publications that use the term *estimation by CBR* describe different type of mechanisms. These include using indexing in retrieval and/or rules in adaptation (Prietula *et al*, 1992; Bisio and Malaboccia, 1995; Prietula *et al*, 1996; McSherry, 1998). Walkerden and Jeffery (1999) have used a linear size adjustment of selected analogies in generating an estimate. This can be perceived as using a rule in generating an estimate. However, they call the process ABE and thus break the pattern otherwise apparent in the literature. This document uses the term ABE to denote the use of simple algorithms for retrieval and adaptation as proposed by Schofield and Shepperd (1995).

Examples of both ABE and estimation by CBR in the construction industry are limited. Hegazy and Moselhi (1994) have created an analogy based solution for markup estimation. However, this involved using artificial neural networks in a domain where problems are usually solved by using analogies. In the construction industry CBR is known to have been used for cost estimation at least in two distinct domains. Marir and Watson (1995) have used it for estimation of refurbishment cost and Perera and Watson (1996) have used it for design and cost estimation of warehouses. In both of these CBR research projects indexing is used in retrieval and rules are used in adaptation. This enforces the pattern related to the use of the terms ABE and CBR.

ANALOGY BASED ESTIMATION PROCESS

As previously noted the ABE process involves the characterisation of the estimated project by attributes and finding analogies from projects used for estimation based on this characterisation. The known target values of the closest analogies are used to generate an estimate for the estimated project (Schofield and Shepperd, 1995).

The process of ABE is fairly straightforward. Schofield and Shepperd (1995) have identified three major considerations in the process:

1. Which attributes to use for the characterisation of the project?
2. How to define the similarity to projects used for estimation?
3. How to use the chosen closest analogies to generate an estimate?

According to Shepperd *et al* (1996) the choice of the estimation attributes to characterise projects remains with the person utilising ABE. The attributes chosen should be the ones that best represent projects in a particular domain.

There are several methods for assessing similarity. The method chosen to be investigated is the Euclidean distance in n-dimensional space as utilised by Schofield and Shepperd (1995). Each of the n dimensions corresponds to a project attribute. The closest analogies are searched by calculating the Euclidean distance to projects used for estimation. The ones with the smallest distances are chosen as analogies. The mathematical expression for Euclidean distance is:

$$E = \sqrt{(a_1\text{-}a_2)^2 + (b_1\text{-}b_2)^2 + ... + (z_1\text{-}z_2)^2} \qquad (1)$$

where E is the Euclidean distance, and

a, b and z are normalised attribute values of projects 1 and 2.

Stensrud and Myrtveit (1998a) criticise the Euclidean distance as a method of identifying closest analogies. They hold the opinion that in order to produce more robust results the input attributes need to have individual weightings. This is because some input parameters are bound to have more effect on the outcome than others. Prior to this criticism Shepperd and Schofield (1997) had proposed three potential solutions to this problem. Firstly, the initial weighting of the input attributes could be done according to expert opinions with subsequent refinement with trial and error. Secondly, the weights could be set according to the relevance of each variable determined by regression analysis. Thirdly, a learning mechanism could be built into ABE that would learn the importance of different attributes and adjust the weights accordingly.

Choosing the number of closest analogies to be used for generating an estimate is problematic. If too few are used, a maverick project may be found. If too many are used, the effects of closest analogies may be masked. Giving the closest analogies a higher weighting has been proposed as a possible solution to this problem (Shepperd and Schofield, 1996). Shepperd *et al* (1996) have found that using the mean of two closest analogies in generating an estimate results in the highest accuracy. They suggest that this is due to the tendency of analogies to over and underestimate.

ANALOGY BASED ESTIMATION IN BUILDING SERVICES

Data Available

The data available for testing ABE for capital cost estimation of building services was twenty costfiles on mechanical building services. The costfiles contained detailed capital cost information, which was broken down into sub-projects and subsequently into components. However, the break down into sub-projects was not consistent with each costfile having its individual classification. Some projects were broken down by area of building e.g. ground floor, first floor or block A, block B etc. Considering the vast amount of detail in the costfiles it was not viable to rearrange the data into a uniform format. Therefore, a decision was taken to test the estimation performance of ABE on complete projects.

The twenty costfiles included ten different types of buildings as illustrated in Table 1.

Table 1. Different building types in the data available.

Building Type	**Number of Buildings**
Factory	3
Hospital	1
Laboratory	2
Office	4
Police Building	1
Recreational Building	2
Retail Building	2
Theatre	1
University Building	3
Warehouse	1

The quantities of various maintenance significant components (MSCs) for each project were extracted from the costfiles. The MSCs have a central role in the overall framework for whole life costing of building services. In ABE the quantities of MSCs are used to characterise each of the projects. The concept of MSC is based on research on cost significant estimation (Poh and Horner, 1995; Horner and Zakieh, 1996; Horenr *et al.*; Al-Hajj and Horner, 1998). The MSCs are the 20 percent of all components that are perceived to constitute 80 percent of the maintenance cost. In the absence of reliable maintenance cost data, expert opinions were used to identify the MSCs. The number of MSCs identified for mechanical building services was 35.

The capital cost for each project was extracted from the costfiles. The capital costs were all normalised to the third quarter of 2000 using the price indices for mechanical services obtained from David Langdon and Everest (2000). It must be noted that some of the cost indices were provisional and some were forecasts. A further price adjustment was carried out according to the location of the project. All costs were converted to a single location using price indices from the same source.

Test Procedure

Two ABE processes were tested independently and subsequently in combination. Firstly, the quantities of MSCs were estimated based on gross floor area and building type. Secondly, the capital cost of complete building services was estimated based on the quantities of MSCs from the costfiles. Thirdly, these two processes were combined and the capital cost was estimated based on gross floor area and building type.

In testing ABE, the jack-knifing method was mainly used. Jack-knifing involves successively removing each project and using the remaining projects for estimation. There were 20 costfiles available for testing. Thus, the ABE process was repeated 20 times. Linear regression (LR) was used for prediction in order to compare the estimation accuracy of ABE. The jack-knifing method was also used in prediction with LR. The test procedure is illustrated in Figure 1.

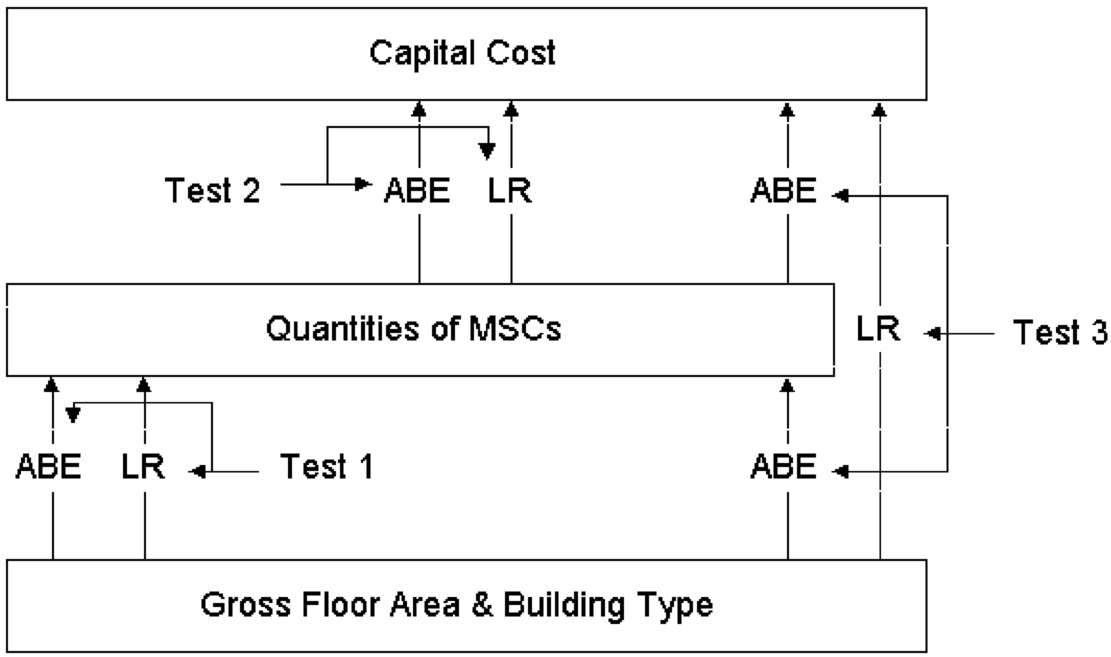

Figure 1. Test procedure.

Test 1 – Estimation of MSCs Using Gross Floor Area & Building Type

Initial estimates of building services are needed at a stage where the information available on the design is incomplete. Using ABE to estimate the quantities of MSCs seeks to generate more information for estimation purposes. Quantities of MSCs were estimated based on gross floor area and building type.

For all the building types presented in the costfiles a mean capital cost per gross floor area ratio was calculated. This is a measure of the intensity of mechanical services. The building types were ranked according to this cost. The least intensive building type was given a rank 1 and the most intensive was given a rank 10. The ranking is illustrated in Table 2.

Table 2. The ranking of building types by intensity of mechanical services.

Building Type	£/m^2	Rank
Warehouse	93	1
Police Building	112	2
Retail Building	131	3
Office	172	4
Factory	172	5
University Building	174	6
Hospital	239	7
Recreational Building	261	8
Theatre	267	9
Laboratory	326	10

The two estimation parameters of gross floor area and the rank of mechanical services intensity were normalised to 1.0. This was done to achieve equal weighting of both parameters on the outcome.

The Euclidean distance was used to measure similarity. The Euclidean distance of the project being assessed to each of the other remaining projects was calculated. The four most similar projects were chosen as analogies to generate estimates. It was anticipated that using more than four closest analogies out of nineteen would mask the effect of closest analogies. An estimate of the quantity of each MSC was generated separately. All the estimates were rounded to the nearest integer. The four closest analogies were used in nine different ways to generate an estimate. These were:

ABE1 the quantity in the closest analogy
ABE2 the mean quantity of the two closest analogies
ABE3 the mean quantity of the three closest analogies
ABE4 the mean quantity of the four closest analogies
ABE5 the mean quantity of the two closest analogies giving the closest a double weighing
ABE6 the mean quantity of the three closest analogies giving the closest a double weighing
ABE7 the mean quantity of the four closest analogies giving the closest a double weighing
ABE8 the mean quantity of the three closest analogies giving the closest a triple and the second closest a double weighting
ABE9 the mean quantity of the four closest analogies giving the closest a quadruple, the second closest a triple and the third closest a double weighting

The estimation accuracy was measured by the magnitude of relative error (MRE) and the mean magnitude of relative error (MMRE). The mathematical expressions for MRE and MMRE are:

$$MRE = \frac{|a_a - a_e|}{a_a} \quad (2)$$

where a_a is a value of an actual attribute of a project, and
a_e is a value of an estimated attribute of a project.

$$MMRE = \frac{\sum_{i=1}^{i=n} \left(\frac{|a_a - a_e|}{a_a} \right)_i}{n} \quad (3)$$

where n is the total number of estimates.

However, MRE and MMRE could not be used to assess the estimation accuracy for the quantities of MSCs. This was because the majority of the target estimates were zero, which occurred when a building did include a specific MSC. This would have led to division by zero in calculating MRE. Therefore, the different approaches to generate estimates were compared in terms of number of closest estimates. Walkerden and Jeffery (1999) use this method as a supplementary comparison of estimation performance.

A quantity for each of the MSC of each project in the costfiles was estimated. The performance of different approaches is illustrated in Table 3. The total number of closest estimates does not equal to the total number of estimates generated. This is because if two or more estimates were equally close to the target quantity then they were all awarded the closest estimate.

Table 3. Number of closest estimates of different approaches to generate estimates.

Approach Used	**Number of Closest Estimates**
ABE1	285
ABE2	303
ABE3	289
ABE4	289
ABE5	267
ABE6	297
ABE7	293
ABE8	295
ABE9	284

It can be observed from Table 3 that the approach using the mean quantity of two closest analogies (ABE2) yielded the largest number of closest estimates. However, it must be noted that the differences in the performances of the nine approaches are small.

Table 4. Number of closest estimates generated by ABE and LR.

Technique Used	Number of Closest Estimates
ABE2	492
LR	483

Subsequently, ABE using the mean quantity of two closest analogies to generate an estimate was compared with the prediction performance of LR. The number of closest estimates was used as a method of assessing estimation performance of ABE and LR. The LR estimates were generated using the FORECAST function of Microsoft Excel. The LR prediction was not fine-tuned e.g. no outliers were removed form the data. This was done to achieve comparability. It must be noted that LR produced some negative estimates of the quantities of MSCs. This indicates that in this problem domain LR can not be used for prediction based on values that are outside the range of values that were used for development of the LR function. The estimation performance of ABE2 and LR is illustrated in Table 4, which indicates that ABE marginally outperforms LR.

Test 2 – Estimation of Capital Cost Using Quantities of MSCs

In testing ABE for capital cost, the input attributes were the quantities of MSCs. The quantities of all components were normalised to 1.0 in order to achieve equal weighting on the outcome. In the process of estimation the project under consideration was calculated Euclidean distances to all projects. The distances were calculated in 35 dimensional space. Each dimension in this space corresponded to a MSC. The four most similar projects i.e. projects with the smallest Euclidean distances were chosen as analogies for generating estimates. Nine estimates were generated following the principles used in Test 1. The principles are presented in Section 4.3.

Estimates using LR were generated for comparison. The jack-knifing method was used in prediction with LR. Therefore, there were only nineteen projects remaining in Costfiles to be used for prediction. This prevented creating a LR function beyond one dependent and seventeen independent variables. Therefore, a regression analysis was run separately for each of the 35 independent variables. Eighteen independent variables with the lowest coefficients of determination were excluded. The coefficient of determination describes the capability of the independent variable to predict the dependent variable. Subsequently, LR analysis was run for the seventeen independent variables to create a prediction function. The function was used to generate an estimate of the capital cost. The LR function could not be established for six datasets. This was due to linear dependencies among the independent variables, which prevented the execution of required calculations.

The mean magnitude of relative error (MRE) was calculated for all ABE and LR estimates. The mean magnitude of relative error (MMRE) was calculated for all the ten approaches to generate estimates including all estimates generated with a particular approach. The estimation accuracy of the ten approaches is illustrated in Table 5.

Table 5. Capital cost estimation accuracy of ABE and LR.

Approach Used	MMRE
ABE1	90%
ABE2	74%
ABE3	73%
ABE4	68%
ABE5	79%
ABE6	75%
ABE7	71%
ABE8	75%
ABE9	71%
LR	452%

It can be observed from Table 5 that using the mean capital cost of the four closest analogies in generating an estimate yielded the most accurate results. The best accuracy obtained with ABE on this set of data was 68%. This accuracy is not sufficient for ABE to be used in industry without further testing. The prediction accuracy of LR at 452% was very poor. Some estimates of capital cost generated with LR were negative values. This confirms that in this problem domain LR cannot be used to predict with data that has not been used in the development of the LR function.

Test 3 – Estimation of Capital Cost Using Gross Floor Area & Building Type

The test of ABE for capital cost based on gross floor area and building type was executed in two stages. Firstly, the quantities of MSCs were estimated with ABE. Subsequently, this information was used to estimate the capital costs. The approaches used in generating estimates from closest analogies were the mean quantity of the two closest analogies in the first stage and the mean capital cost of four closest analogies in the second stage. These were the approaches that were found to have the superior performance. - see sections 4.3 and 4.4. The jack-knifing method was used in the first stage, but not in the second stage. This is because none of the sets of quantities of MSCs estimated in the first stage featured in the data used for estimation in the second stage.

ABE was compared with the prediction performance of LR. The estimates using LR were generated based on gross floor area. It was not practicable to develop separate functions for all building types due to the largest number of costfiles on any one building type being four. MRE and MMRE were calculated for the ABE and LR estimates. The performance of ABE and LR is illustrated in Table 6.

Table 6. Overall estimation accuracy of ABE and LR.

Approach Used	MMRE
ABE	76%
LR	55%

It is shown in Table 6 that LR outperforms ABE when estimation is carried out based on gross floor area and building type. However, the estimation accuracy of neither technique is sufficient for them to be used in industry. The poor estimation accuracy obtained by ABE is anticipated to be partially due to the lack and inconsistency of data. There is an obvious shortage of appropriate analogies as ten different building types are represented in dataset of twenty projects.

FURTHER DEVELOPMENT

It is quite obvious that an analogy based estimation can not be used for capital estimation of building services in industry based on the data used in this study. The estimation performance of ABE was poor, even if it outperformed LR on two of the three test settings. However, the tests described in this document can only be seen as a starting point for the development and testing of ABE in building services. No conclusions can be made of the applicability of the technique in the problem domain at this stage. However, scope for the development of ABE in building services can be readily identified.

The data obtained for the tests described in this document was inadequate and inconsistent. This highlights the need for systematic data collection. The value of high quality data can not be emphasised enough. A meaningful estimation tool can only be developed if a sufficient amount of quality data is available. The output of any estimation tool will be only as good as the data used as input.

It is anticipated that the performance of ABE would improve if a separate estimate were generated for each building type and building services system. This is perceived to increase the probability of selecting more appropriate analogies for generating estimates, as there would be less variation in the past solution. Several buildings services system classifications exist. BISRIA is currently promoting a classification by Nanayakkara and Fitzsimmons (1999) as a cost benchmarking standard for capital costs of building services. High quality data is required to test this proposition.

Using ABE to estimate quantities of MSCs has an inherent weakness. The estimated quantities may not be consistent with the design principles of building services. For example, air conditioning can be provided for in different ways like variable refrigerant flow or chilled ceilings and beams. Using ABE to estimate the quantities of MSCs may include components from both ways of providing air conditioning. Hence, there is need to ensure that the estimates generated comply with the design principles. This could be achieved by using rules in generating estimates. However, then in could be argued that the technique should be called CBR rather than ABE.

Stensrud and Myrtveit (1998a) and Stensrud and Myrtveit (1998b) have studied the estimation performance of ABE, LR and expert estimators. They have demonstrated that expert judgement improves the estimation performance of ABE more than it does that of LR. They suggest that this is because recognising when an estimate is to be trusted and when disregarded as misleading requires expert judgement, which can be done more readily with ABE than LR. According to Stensrud and Myrtveit (1998a), there is more value in ABE in identifying the closest analogies than in generating the estimates. They subsequently argue that the ultimate test of an estimation tool is to evaluate how much it improves human performance. In their opinion the performance of the estimation tool itself is of secondary interest (Stensrud and Myrtveit, 1998b). In the light of these arguments the performance of ABE in building services should also be evaluated in combination with expert judgement.

REFERENCES

Aamodt, A. and Plaza, E. (1994) Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. *AI Communications*, **7** (i), 39-59.

Al-Hajj, A. and Horner, M. (1998) Modelling the Running Costs of Buildings. *Construction Management and Economics*, **16** (4), 459-470.

Bisio, R. and Malabocchia, F. (1995). Cost Estimation of Software Projects Through Case Based Reasoning. *Proceedings of the International Conference on Case Based Reasoning,* Sesimbra, Portugal.

Cowderoy, A. and Jenkins, J. (1988) Cost Estimation by Analogy as a Good Management Practice. *Proceedings of Software Engineering*, Liverpool, England, 80 – 84.

Davis Langdon and Everest (2000) *Spon's Mechanical and Electrical Services Price Book 2001*. E&FN Spon, London.

Hegazy, T. and Moselhi, O. (1994). Analogy Based Solution to Markup Estimation Problem. *Journal of Computing in Civil Engineering,* **8** (1), 72-87.

Horner, R., El-Haram, M. and Munns, A. (1997) Building Maintenance Strategy: A New Management Approach. *Journal of Quality in Maintenance Engineering*, **3** (4), 273-280.

Horner, R. and Zakieh, R. (1996) Characteristic Items – A New Approach to Pricing and Controlling Construction Projects. *Construction Management and Economics,* **14** (3), 241-252.

Mair, C., Kadoda, G., Lefley, M., Phalp, K., Schofield , C., Shepperd, M. and Webster, S. (2000) An Investigation of Machine Learning Based Prediction Systems. Forthcoming in the *Journal of Systems and Software*.

Marir, F. and Watson, I. (1995) CBRefurb: A Case-Based Building Refurbishment Cost Estimator and Decision Support System. Topping, B. (ed.) *Developments in Artificial Intelligence for Civil and Structural Engineering*, 231-236.

McSherry, D. (1998) An Adaptation Heuristic for Case-Based Estimation. *Proceedings of 4th European Workshop, Advances in Case-Based Reasoning*, Berlin, Germany, 184 – 195.

Nanayakkara, R. and Fitzsimmons, J. (1999) *Cost Benchmarks for the Installation of Building Services: Part 3 - A Protocol for Recording Project Costs.* BSRIA, Bracknell.

Perera, S. and Watson, I (1996) Multi-Agent Collaborative Estimating and Design in NIRMANI: Organising a Multi-Perspective Case Memory. Kumar, B. (ed.) *Information Processing in Civil and Structural Engineering Design*, 53-64.

Prietula, M., Vincinanza, S. and Mukhopadhyay, T. (1992) Extraction and Representation of Software Estimation Expertise in a Case-Based Reasoning. *Proceedings of International Symposium on Artificial Intelligence,* Menlo Park, CA, USA, 232-237.

Prietula, M., Vincinanza, S. and Mukhopadhyay, T. (1996) Software Effort Estimation With a Case-Based Reasoner. *Journal of Experimental & Theoretical Artificial Intelligence*, **8**, 341 - 363.

Poh, P. and Horner, R. (1995) Cost-Significant Modelling – Its Potential for Use in South-East Asia. *Engineering, Construction and Architectural Management,* **2** (2) 121-139.

Reisbeck, C and Schank, R. (1989) *Inside Case-Based Reasoning.* Lawrence Erlbaum Associates, Hillsdale, USA.

Ritter, R. (1980) Logic for Default Reasoning. *Artificial Intelligence*, **13** (March).

Schofield, C. and Shepperd, M. (1995) Software Support for Cost Estimation by Analogy. *Proceedings of 6th European Software Cost Modelling Conference*, Rolduc, Netherlands.

Shepperd, M. and Schofield, C. (1996) Effort Estimation by Analogy: A Case Study. *Proceedings of the 7th European Software Control and Metrics Conference*, Wilmslow, England.

Shepperd, M. and Schofield, C. (1997). Estimating Software Project Effort Using Analogies. *IEEE Transactions on Software Engineering*, **23** (11), 736-743.

Shepperd, M., Schofield, C. and Kitchenham, B. (1996) Effort Estimation Using Analogy. *Proceedings of the 18th International Conference on Software Engineering*, Berlin, Germany, 170 – 178.

Stensrud, E. and Myrtveit, I. (1998a) The Added Value of Estimation by Analogy: An Industrial Experiment. *Proceedings of the European Software Measurement Conference*, Antwerpen, Germany, 549-56.

Stensrud, E. and Myrtveit, I. (1998b) Human Performance Estimating with Analogy and Regression Models: An Empirical Validation. *Proceedings of the 5th International Software Metrics Symposium*, Bethesda, Maryland, USA, 205 - 212.

Walkerden, F. and Jeffery, R. (1999) An Empirical Study of Analogy-Based Software Effort Estimation. *Empirical Software Engineering*, **4**, 135-158.