Engineering, Construction and Architectural Management 2001 8 | **3**, 198–210

Addressing the contractor selection problem using an evidential reasoning approach

M. SÖNMEZ*, J. B. YANG* & G. D. HOLT†

**Manchester School of Management, UMIST, PO Box 88, Manchester M60 1QD and †The Built Environment Research Unit, University of Wolverhampton, West Midlands WV1 1SB, UK*

Abstract Selecting the 'best' main contractor is a complex decision process for construction clients. It requires a large number of criteria to be simultaneously measured and evaluated. Many of these criteria are related to one another in a complex way and therefore, they very often conflict insofar as improvement in one often results in decline of another(s). Furthermore, as contractors' attributes are expressed in both quantitative and qualitative terms, decision-makers have to base their judgements on both quantitative data and experiential subjective assessments. In this paper, the evidential reasoning (ER) approach (which is capable of processing both quantitative and qualitative measures) is applied as a means of solving the contractor selection problem (CSP). The process of building a multiple criteria decision model of a hierarchical structure is presented, in which both quantitative and qualitative information is represented in a unified manner. The CSP is then fully investigated using the ER approach. Both the advantages of applying this model in practice and the analysis process itself are discussed.

Keywords bidder evaluation, contractor selection, evidential reasoning, linguistic variables, multiple criteria decision analysis, preference elicitation

INTRODUCTION

In increasingly competitive global markets, accurate and efficient decision-making has become more important than ever before. This is because, once a decision is made the resources (e.g. labour, materials, capital) consumed in achieving it may not be recoverable; whether the decision is robust or not. Decision-makers (DMs), therefore, very often need to think hard, and devote much time and effort to such business 'problems'. This is even more so where subjective (decision) criteria have to be taken into account. In such cases, it would be helpful if a systematic procedure were available to deal with this subjective decision making complexity. Typically, a decision problem is said to be complex and difficult, where the following conditions apply:

- multiple criteria exist, which can be both quantitative and qualitative in nature;
- uncertainty and risk is involved;
- there may be multiple DMs;
- decision (input) data may be vague, incomplete or imprecise (Hipel *et al.*, 1993).

In trying to select the 'best' contractor [hereafter termed the contractor selection problem (CSP)], the task facing a construction client is a multiple criteria decision-making (MCDM) process, in which a large number of criteria need to be evaluated (Hatush & Skitmore, 1998). Most of these criteria are related to each other in a complex way. Furthermore, many usually conflict, such that a gain in one criterion requires a trade-off in another(s). As CSP decision criteria are a mix of both qualitative and quantitative characteristics, DMs have to base their decisions on both quantitative analysis and subjective (typically experiential) judgements. DMs may intuitively find it easier to make subjective judgements by using verbal expressions (i.e. linguistic variables) (Poyhonen *et al.*, 1997). However, this can cause problems during evaluation of alternatives, because it is difficult to process (e.g. aggregate) these two types of measure (i.e. quantitative and linguistic). It is, therefore, necessary that any MCDM method be capable of aggregating these two types of measures in a rational and consistent manner; ultimately providing a ranking of all decision alternatives. The evidential reasoning (ER) approach was developed on the basis of decision theory and uses the Dempster–Shafer theory of evidence (see later), to aggregate these 'conflicting' types of assessment (Yang & Sen, 1994; Yang, 2001). In this respect, ER shows significant potential to solve the CSP (Holt, 1998). To date, ER has been used mainly in 'artificial intelligence' and 'expert systems' as a technique for modelling reasoning under uncertainty (Beynon *et al.*, 2000). Uniquely, a demonstrative application of ER to the CSP is elucidated in this paper.

In practice, contractor selection is normally a two-stage process whereby contractors are first pre-qualified (for example, to get onto a select list or be invited to tender for a given project). Subsequently, their tender submissions are evaluated in the second stage. This paper describes a method to perform that second stage; this being mainly to simplify the narrative for elucidation of the technique. Nonetheless, the ER approach can accommodate either of the two selection stages mentioned above.

The structure of the paper is as follows. The difficulty of knowledge acquisition is explained following the introduction. Then, the process of building a MCDM model of a hierarchical structure is presented, in which both quantitative and qualitative information is processed in a unified manner (through equivalent knowledge transformation). The ER approach is then explained and applied to the CSP using computer software in the fourth section of the paper. Results and discussion follow with the advantages of the ER approach in this context being reported.

KNOWLEDGE ACQUISITION

Understanding a DM's attitudes towards uncertainty and risk is an important ingredient for effective decision making. During the conventional evaluation process of a decision-making problem, individuals are often required to give exact or precise numerical assessments with regard to each decision criterion. Although this is achievable, the numerical representation of subjective characteristics may impose a heavy burden on the DM. Furthermore, these numbers may not 'truly' reflect the DM's preferences. It is desirable that DMs not be forced to provide exact numerical assessments, but rather, should be free to express their judgements either numerically or subjectively. The ER method facilitates this.

MCDM methods as preference elicitation tools

There are many MCDM methods proposed in the literature (see Hwang & Yoon, 1981, for a classification of these methods), each having different ways of eliciting a DM's assessments in order to evaluate alternatives based on multiple criteria. The analytical hierarchy process (AHP) developed by Saaty, for example, asks DMs to compare alternatives in a pair-wise fashion based on each decision criterion (Saaty & Wind, 1980). Here, DMs are required to make exact or precise statements like 'I think alternative A is three times more important than alternative B as far as a particular criterion is concerned'. However, it may be unrealistic to expect all DMs to be able to provide such kind of statements, because of, for example, the complexity of decision problems, a lack of data and shortcomings in expertise. The AHP method has been criticized by some academics because: (i) of the scale used (Poyhonen *et al.*, 1997), (ii) it requires redundant information from the DM (Islei & Lockett, 1988), (iii) the occurrence of rank reversals and (iv) the comparison of two criteria represented by two totally different scales (Belton & Gear, 1983, 1985; Belton, 1986; Stewart, 1992).

Multiple attribute utility theory (MAUT) on the other hand, uses the concept of utility to determine a DM's real preferences, judgements and attitudes towards risk (Keeney & Raiffa, 1993). However, this approach also places a burden on DMs by asking a large number of hypothetical lottery-type questions in order to discover their real preferences. Subsequently, DMs may not give consistent answers to these questions. Like the AHP, MAUT requires DMs to provide exact numbers (i.e. probability values) so that their utility functions can be derived. Another disadvantage of MAUT is that the decision-making process takes a long time and becomes tedious if there are numerous criteria. The method also to some extent pre-supposes that DMs are very good at probability theory, which may not be the case in reality. Hatush & Skitmore (1998) used MAUT to solve a CSP. The aim of the present paper uses a similar CSP as an example, to show how the ER approach can be conveniently, and effectively, applied to deal with this kind of problem.

Literature review on contractor selection

The recent literature on contractor selection can be divided into three groups: (i) CSP criteria (attributes) and their weights, (ii) criteria measurement and (iii) selection methodologies. Several academics have studied the decision criteria used by clients for choosing a contractor (Russell & Skibniewski, 1988; Holt *et al.*, 1994a; Ng, 1996; Hatush & Skitmore, 1997a). Holt *et al.* (1994a) carried out a survey of 53 major UK construction client organizations to determine the decision criteria used for contractor selection and the importance of these criteria in terms of influencing their choice of contractors. Hatush & Skitmore (1997a) found that all clients use a 'similar' set of criteria for contractor selection, but that the way clients quantify these criteria can be very different in practice. In these previous works, a contractor's bid amount appears to be the most dominant and important criterion (Holt *et al.*, 1993, 1994c; Hatush & Skitmore, 1997a, 1998). However, choosing a contractor based solely on the lowest bid price is one of the major causes of project delivery problems. Another disadvantage of using the lowest bid as a principal discriminating criterion is that some contractors (e.g. facing a shortage of work) may enter unrealistically low bid prices, simply to try and maintain cashflow. Therefore, as Hatush & Skitmore (1997a) indicated, financial and technical criteria must be considered in order to assess the potential of contractors finishing projects on time; and to assess whether contractors have the necessary resources to complete any contract awarded to them.

The following four weaknesses were found in contractor selection practice: (i) lack of a universal approach, (ii) long-term confidence attributed to results of pre-qualification, (iii) reliance on tender sum in decision making and (iv) inherent subjectivity of the process (Holt *et al.*, 1993, 1995a). A summary of current evaluation strategies can be found in Hatush & Skitmore (1997a,b). Holt *et al.* (1994b) classified the contractor selection process into three stages: (i) pre-qualification, (ii) contractor evaluation and (iii) final selection. For each stage, three types of scores were proposed (P1, P2 and P3, respectively). P1 scores represent the general organizational attributes of a contractor and also provide insight of specific contractor weakness(es). A multiattribute analysis (MAA) technique was used to combine P2 scores (representing the scores of project-specific criteria) and P3 scores (representing bid amount) into a simple index. This index was determined by assigning a 40% weighting to the P2 scores and a 60% weighting to the P3 scores (sensitivity analysis revealed these percentages to the best discriminate among contractors).

Holt *et al.* (1993, 1994a,b,c,d, 1995b) provided example application of MAA to the evaluation of construction bidders. They developed a method to evaluate contractor pre-qualification criteria and provided guidelines for practitioners, highlighting areas to address when evaluating a contractor based on a particular criterion (Holt *et al.*, 1994d). Holt *et al.* (1996) applied cluster analysis as a means of reducing a large number of potential bidders, to identify only those suitable to tender for a particular project. Ng (1996) investigated different decision support systems (DSS) for contractor pre-qualification. Amongst the surveyed DSS's were database management systems (DBMS), expert systems (ES), fuzzy sets (FS) and case based reasoning (CBR). Database management systems is a proprietary software capable of storing and retrieving data with the help of a user friendly query language [i.e. structured query language (SQL)] (Kerry, 1990). The contractor management information system (CMIS) maintained by the Department of the Environment is probably the most universally recognized system in the UK and stores around 8000 records of contractors who work for public clients (Deparment of Environment, 1992). Expert systems mimic the problem solving process of users in a particular problem domain (Adelman, 1992). The two types of ES – rule-based and object oriented – were discussed in Ng (1996). Russell *et al.* (1990) developed a rule-based ES called 'QUALIFIER-2' for contractor pre-qualification. Ng (1993) developed an integrated object oriented ES for contractor pre-qualification, whilst Taha *et al.* (1995) proposed a knowledge-based DSS for predicting construction contract bond claims using contractor financial data. This DSS employed inductive learning and neural networks to extract the problem solving knowledge. The concept of FS was first introduced by Zadeh (1965) to deal with fuzzy and uncertain data that is typically represented by linguistic, rather than numeric, variables. Nguyen (1985) proposed and applied an FS model to contractor pre-qualification and tender evaluation. CBR reuses or modifies experiential knowledge to solve problems, which are ill defined and contain both qualitative and quantitative criteria. Hence, Ng (1996) found CBR as a suitable tool to study contractor pre-qualification.

Hatush & Skitmore (1997b) applied programme evaluation and review technique (PERT) to assess and evaluate contractor data against client goals (time, cost and quality). Hatush & Skitmore (1998) used MAUT to select the best contractor based on a mixture of qualitative and quantitative criteria. In a recent study, Holt (1998) reviewed the use of different CSP methods and the following were identified as having been applied in this context: bespoke approaches, MAA, MAUT, cluster analysis, multiple regression, fuzzy set theory and multivariate discriminant analysis. The advantages and disadvantages of these methods were also discussed. Despite this previous research, the problem of reconciling quantitative and qualitative CSP data remains. It is this aspect that the present paper concentrates upon by applying the ER technique.

THE ER APPROACH TO CSP

One mistake that DMs often make is to try to solve a decision-making problem straight away. As a result of this, because of lack of forethought, very often the problem is approached incorrectly. Therefore, DMs should understand and have a clear picture of the whole problem before they start trying to solve it. This is much more important when there are many criteria to consider, which in turn may comprise sub criteria and even sub, sub criteria. For this reason, it is useful to display the problem in the form of a hierarchical structure, as follows.

The DM is to choose an alternative a_i from a finite number of alternatives $a_1, a_2, a_3, \ldots, a_n$ ($i = 1, 2, 3,\ldots, n$). These alternatives have to be evaluated based on m main criteria $c_1, c_2, c_3,\ldots, c_m$. Each main criterion may have a different number of k sub criteria such that $c_{i1}, c_{i2}, \ldots, c_{ik}$ ($i = 1, 2, 3, \ldots, m$). It is necessary to assign weights to the main criteria according to their contribution to the overall objective $w_1, w_2, w_3, \ldots, w_j$ ($j = 1, 2, 3, \ldots, m$) and also to the sub-criteria $w_{i1}, w_{i2}, w_{i3}, \ldots, w_{ik}$ ($i = 1, 2, 3,\ldots, m$). This is so as to show the relative importance of each sub-criterion to its associated upper level criterion. Several methods for weight assignment have been proposed in the literature (Barron & Barrett, 1996; Sen & Yang, 1998, pp. 26–43). These weights are used for propagating lower level criteria assessments to respective upper levels. For simplicity, in this paper the same set of decision criteria and weights are used as proposed in Hatush & Skitmore (1998). These decision criteria and their weights are shown in Fig. 1.

The ER approach

ER has increasingly been used in a diverse range of areas ranging from engineering, management, to safety and has been applied to different MCDM problems. Interested readers may refer to the following references for a full explanation of the method and its associated algorithm: Yang & Singh (1994), Yang & Sen (1994, 1997), Wang *et al.* (1995, 1996), Yang (2001). Recently, Beynon *et al.* (2000) gave a number of simple examples to explain the Dempster–Shafer theory of evidence (DST). The ER approach uses the concept of 'degree of belief (DoB)' as a preference elicitation tool. The DoB can be described as the degree of expectation that an alternative will yield an anticipated outcome on a particular criterion. An individual's DoB depends on their knowledge of the subject and their experience. The use of the DoB can be justified by the fact that human decision making involves ambiguity, uncertainty and imprecision. That is, individuals can convey judgements in probabilistic terms with the help of their knowledge and real life experience. Probability has long been used to deal with uncertainty and risk in decision problems; it can be a powerful tool to overcome the imprecision and ambiguity of human decision making.

Decision problems are usually structured in a hierarchical order (refer Fig. 1). In the first level, the goal of the problem is stated. In the second level, there are several criteria, each of which has a different contribution to measuring, and helping achieve the overall goal. Then, some of these criteria may be broken down into further sub-criteria. The process (i.e. disaggregating main criteria into sub-criteria, and then sub-criteria into sub, sub-criteria) continues up to the point where DMs are able to make practical assessments (on these lower level criteria). Once the subdivision of criteria is complete, DMs evaluate each alternative based on the lowest level criteria. In order to find out how well an alternative performs across all criteria, the lowest level criteria assessments need to be first transformed to their relevant upper levels and ultimately, to the top-level goal. This requires an appropriate MCDM method. The ER approach is such a method that cannot only combine both qualitative and quantitative assessments, but can also handle uncertain and imprecise information or data.

Implementation of the ER approach

The ER approach can be described as a hierarchical evaluation process in which all decision criteria are aggregated into one (i.e. the goal of the problem). As the ER algorithm has previously been well-explained (Yang & Sen, 1994; Yang, 2001), the ER process is briefly described here in a stepwise manner:

1. display a decision problem in a hierarchical structure;
2. assign weights to each (main) problem criterion and also to their sub-criteria (if any);
3. choose a method for assessing a criterion either quantitatively or qualitatively;
4. transform assessments between a main criterion and its associated sub-criteria if they are assessed using different methods (i.e. quantitative and qualitative);
5. evaluate each alternative based on the lowest (i.e. bottom) level criteria in the hierarchical structure;
6. quantify qualitative assessments at the top level if necessary and determine an aggregated value for each alternative;
7. rank alternatives based on this aggregated value and (normally) choose the highest rank.

The ER algorithm is integrated into a software package called 'intelligent decision system via evidential reasoning' (IDS[1]1A demonstration version of IDS with the example of CSP presented in this paper can be obtained from Dr J.B. Yang through e-mail request: jian-bo.yang@umist.ac.uk.). Intelligent decision system

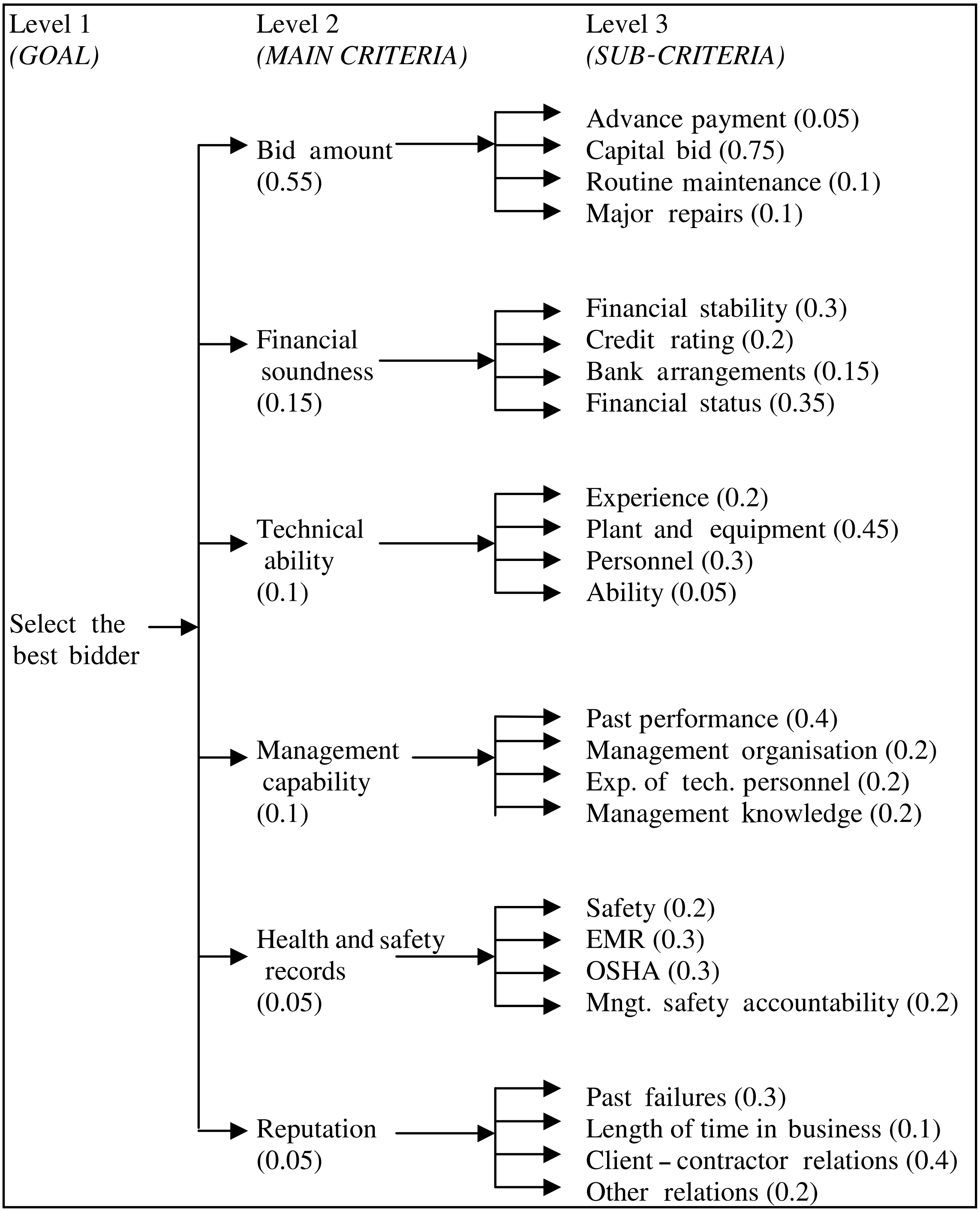

Figure 1 Hierarchical display of the CSP and the relative importance of criteria and sub criteria (Source: Hatush & Skitmore, 1998).

is a Windows™ based and graphically designed decision support tool developed by Yang & Xu (2000). It allows DMs to build their own models and input their own data. The CSP is hierarchically displayed and weights are assigned to the decision criteria (Fig. 1). Alternatives (i.e. contractors) can be assessed using clearly defined standards (grades). It is assumed that a single decision-maker (DM) is involved in the decision-making process who is first asked to define assessment grades for the goal of the problem. Let us suppose that the DM wants to classify contractors being evaluated, into the following grades: worst, poor, average, good and excellent at the top level. Next, the DM is required to define assessment grades for the main criteria. The outcomes for each criterion may be expressed in different terms in the mind of the DM who may wish to use the most appropriate vocabulary to evaluate (and represent) each criterion. Therefore, the DM may well use the same set of grades as defined for the goal of the problem for some main criteria, and develop new sets of grades for other main criteria. In Fig. 2, the DM used seven grades for the criterion 'financial soundness' whilst the other main criteria were evaluated with a set of five grades each, using different wordings. The use of different grades facilitates data collection and allows capture of the DM's preferences, experience, intuition or beliefs and also implies that the DM is not manipulated (by the method or decision analyst who may help them during the decision process). This is because they use their own expressions to evaluate decision criteria. Although this may increase ambiguity, uncertainty or imprecision in the data, the ER approach facilitates this through rule and utility based knowledge transformation, which will be explained in the subsequent sections.

In a similar manner to that shown in Fig. 2, the DM is asked to assign classification grades to the bottom level criteria. Note that again different sets of grades were used at the bottom level criteria shown in Fig. 3 in order to evaluate contractors based on the sub-criteria. If a sub-criterion is evaluated quantitatively, then there is no need to define grades for it. For example, because the sub-criteria of the criterion 'bid amount' are all of a quantitative nature, these monetary values are used to evaluate the contractors.

Next, the DM is asked to evaluate each contractor based on the sub-criteria by using the grades defined in Fig. 3. Suppose the evaluation yields the assessments shown in Fig. 4, which are obviously multivariate: (1) It is noted that on some of the criteria, assessments for different contractors are missing. Take contractor K, for example, who has no assessment on the criterion (technical ability) 'personnel'. This may be because of the fact that contractor K has not provided any information on this criterion. (2) Some of the assessments are incomplete and uncertain as the total DoB is less than unity. Again, contractor K has been assessed to be 50% average and 40% strong in terms of financial stability. This assessment is said to be uncertain and incomplete because the total DoB is less than 1 (0.9). The 'missing' 10% may represent the doubts of the DM, a lack of evidence and/or missing information. (3) Some assessments are certain. Contractor K is rated as 100% on the criterion 'management organization'. This assessment is said to be certain and complete because the DoB is equal to 1. In order to make more accurate and complete assessments, it might be useful to provide the DM with a checklist that shows the areas to address and what to look at, as suggested in Holt *et al.* (1994a). This checklist would also be of help in reminding the DM of the grades associated with criteria. For example, when assessing a contractor's past performance, suppose there are five areas to address. The contractor is said to have a 100% very good past performance, if the contractor has a very good record in all these areas.

Assessment transformation

Here, it is shown how the ER approach deals with the multidisciplinary data given in Fig. 4. The process of transforming the bottom level criteria assessments to the associated upper level criterion will be explained. The assessment transformation can be made using either the rules given by the DM or by using utility theory. Take the criterion 'advance payment', for example, it is a sub-criterion of the upper level criterion 'bid amount'. As the bid amount is classified against one of five grades (very high, high, average, low and very low), the DM is required to convert the monetary values into one of these descriptive grades. For this example, it is assumed that the DM expressed these as shown in Fig. 5.

Main criteria	Assessment grades						
Bid amount	Very high	High	Average	Low	Very low		
Financial soundness	Worst	Very bad	Bad	Indifferent	Good	Very good	Excellent
Technical ability	Very low	Low	Average	High	Very high		
Management capability	Very low	Low	Average	High	Very high		
Health and safety records	Very poor	Poor	Indifferent	Good	Very good		
Reputation	Very bad	Bad	Average	High	Very high		

Figure 2 Assessment grades defined by the DM for the second level criteria.

Sub-criteria	Assessment grades						
Advance payment	Quantitative						
Capital bid	Quantitative						
Routine maintenance	Quantitative						
Major repairs	Quantitative						
Financial stability	Very weak	Weak	Average	Strong	Very strong		
Credit rating	Very low	Low	Average	High	Very high		
Bank arrangements	Worst	Very bad	Bad	Normal	Good	Very good	Excellent
Financial status	Very unreliable	Unreliable	Average	Reliable	Very reliable		
Experience	Very inexperienced	Inexperienced	Average	Experienced	Very experienced		
Plant & equipment	Very old	Old	Average	New	Very new		
Personnel	Less trained	Average	Trained	Highly trained			
Ability	Low	Average	High				
Past performance	Very poor	Poor	Average	Good	Very good		
Project management organisation	Poor	Average	Good	Excellent			
Experience of technical personnel	Very unsatisfactory	Unsatisfactory	Average	Satisfactory	Very satisfactory		
Management knowledge	Low	Adequate	High	Excellent			
Safety	Very unsafe	Unsafe	Average	Safe	Very safe		
Experience modification rate (EMR)	Very low	Low	Average	High	Very high		
Occupation safety (OSHA)	Inadequate	Adequate	Good	Excellent			
Management safety accountability	Very low	Low	Average	High	Very high		
Past failures	Quantitative						
Length of time in Business	Quantitative						
Client / contractor relationships	Very poor	Poor	Average	Good	Very good		
Other relations	Very poor	Poor	Average	Good	Very good		

Figure 3 Assessment grades defined for sub-criteria.

According to this information, contractor K is said to have a low bid amount as far as the criterion 'advance payment' is concerned (refer Fig. 4). Contractor N's advance payment is £0.15 m. This amount is converted to 50% average and 50% low because it is mid-way between these two grades (refer Fig. 5). When a contractor is evaluated based on the criterion 'financial soundness', for example, the following sub-criteria (attributes), such as financial stability, credit rating, bank arrangements and financial status could be used (Fig. 1). These sub attributes are assessed by subjective judgements (Fig. 3). Because a different number of grades are used for the upper level criterion and the sub-criteria, the DM is asked to establish rules to propagate sub criteria assessments to the associated upper level criterion. The rules given by the DM are shown in Fig. 6.

		Contractors				
Criteria	Sub-criteria	K	L	M	N	O
Bid amount	Advance payment (m£)	0.1	0.3	0.3	0.15	0.1
	Capital bid (m£)	3.9	3.5	3.5	4.0	3.6
	Routine (m£) maintenance	0.3	0.25	0.3	0.25	0.1
	Major repairs (m£)	0.4	0.35	0.2	0.4	0.4
Financial soundness	Financial stability	A(0.5) S(0.4)	A(0.75) S(0.2)	A(0.4) S(0.5)	A(1.0)	A(0.9)
	Credit rating	A(0.30) H(0.60)	A(0.15) H(0.80)	A(0.40) H(0.55)	No information	A(0.80) H(0.20)
	Bank arrangements	N(0.25) G(0.75)	N(0.75) G(0.25)	N(0.25) G(0.75)	B(0.33) N(0.67)	N(0.75) G(0.25)
	Financial status	A(0.17) R(0.83)	A(0.17) R(0.83)	A(0.33) R(0.67)	U(0.17) A(0.83)	A(0.67) R(0.33)
Technical ability	Experience	I(0.10) A(0.80)	A(0.20) E(0.75)	I(0.40) A(0.50)	E(1.0)	I(0.80) A(0.05)
	Plant and equipment	A(0.60) N(0.30)	A(0.25) N(0.55)	O(0.5) A(0.5)	N(0.4) VN(0.4)	N(0.65) VN(0.2)
	Personnel	No information	A(0.2) T(0.8)	A(0.2) T(0.6)	T(1.0)	LT(0.6) A(0.2)
	Ability	L(0.17) A(0.83)	No information	A(0.63) H(0.37)	A(0.87) H(0.13)	L(1.0)
Management capability	Past performance	G(0.83) VG(0.17)	A(1.0)	G(0.67) VG(0.33)	A(1.0)	A(1.0)
	Management organisation	A(1.0)	G(0.75) E(0.25)	A(0.5) G(0.5)	A(0.85)	A(0.83) G(0.17)
	Experience of tec. Personnel	A(1.0)	S(1.0)	U(0.25) A(0.75)	U(0.75) A(0.25)	A(0.5) S(0.5)
	Management knowledge	H(1.0)	H(1.0)	A(0.2) H(0.8)	H(0.2) E(0.8)	No information
Health and safety records	Safety	U(0.3) I(0.5)	S(0.55) VS(0.3)	S(0.7) VS(0.2)	U(0.3) I(0.6)	S(0.45) VS(0.4)
	EMR	A(0.4) H(0.6)	L(1.0)	No information	VL(0.4) L(0.6)	VH(1.0)
	OSHA	A(0.86) G(0.14)	A(0.14) G(0.86)	A(0.71) G(0.29)	A(0.57) G(0.43)	G(0.67) E(0.33)
	Management safety	L(0.8) A(0.2)	A(1.0)	A(0.75) H(0.25)	L(0.6) A(0.4)	A(1.0)
Reputation	No. of Past failures	5	4	9	10	9
	Length of time in business	42	45	42	33	18
	Client–contractor relationship	P(0.5) A(0.5)	A(0.75) G(0.25)	A(0.5) G(0.5)	No information	P(0.4) A(0.5)
	Other relations	P(0.65) A(0.25)	A(1.0)	G(0.55) VG(0.25)	P(0.75) A(0.25)	A(0.75) G(0.25)

Figure 4 The assessment of contractors based on the six main criteria (note that the assessment grades are abbreviated).

In Fig. 6, the DM indicated that a contractor with *strong* financial stability means that the financial soundness of this contractor is 85% *good* and 15% *very good* as far as financial stability is concerned. As far as selecting the best bidder is concerned, a contractor with a *good* financial soundness is considered to be 20% *average*

Bid amount	Very high	High	Average	Low	Very low
Advance payment (£m)	0.4	0.3	0.2	0.1	0.05

Figure 5 Transforming a quantitative sub-criterion assessment to the associated upper level criterion.

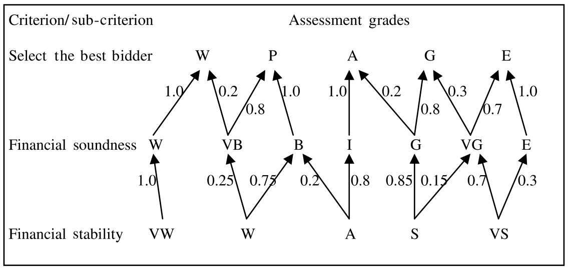

Figure 6 The process of converting lower level criterion assessments to the upper level criterion.

and 80% *good*. A contractor with a *very good* financial soundness is, on the other hand, considered to be 30% *good* and 70% *excellent*. Such rules are based on evidence and the past experience of the DM. As required by the ER approach, individuals are asked to describe the assessment grades so that results can be interpreted, be quantified and be reproducible. Rules for other criteria can be provided in a similar way. If the DM provides no rule, the ER approach assumes that the DM is neutral to risk and the utility of the grades are equidistantly assigned in the normalized space.

RESULTS AND DISCUSSION

The assessments given by the DM in Fig. 5 are fed into IDS and the aggregated results are yielded at the main criteria level (refer Table 1). The assessment grades for each main criterion are abbreviated in Table 1. The numbers in brackets show the degrees of belief of the DM that are aggregated from the assessments of the sub-criteria. One can rank the contractors for each criterion in order of preference by comparing the distributed assessments shown in

Table 1 Combined assessments of the contractors at the main criteria level.

	Contractors				
Main criteria	K	L	M	N	O
Bid amount	VH (0.579)	VH (0.000)	VH (0.000)	VH (0.915)	VH (0.000)
	H (0.390)	H (0.029)	H (0.032)	H (0.034)	H (0.516)
	A (0.016)	A (0.971)	A (0.936)	A (0.043)	A (0.436)
	L (0.015)	L (0.000)	L (0.032)	L (0.008)	L (0.032)
	VL (0.000)	VL (0.000)	VL (0.000)	VL (0.000)	VL (0.016)
Financial soundness	W (0.000)	W (0.000)	W (0.000)	W (0.000)	W (0.000)
	VB (0.000)	VB (0.000)	VB (0.000)	VB (0.005)	VB (0.000)
	B (0.029)	B (0.041)	B (0.026)	B (0.139)	B (0.054)
	I (0.243)	I (0.363)	I (0.295)	I (0.630)	I (0.732)
	G (0.593)	G (0.487)	G (0.563)	G (0.025)	G (0.160)
	VG (0.085)	VG (0.084)	VG (0.077)	VG (0.000)	VG (0.024)
	E (0.000)	E (0.000)	E (0.000)	E (0.000)	E (0.000)
Technical ability	VL (0.001)	VL (0.000)	VL (0.000)	VL (0.000)	VL (0.105)
	L (0.021)	L (0.015)	L (0.333)	L (0.003)	L (0.267)
	A (0.478)	A (0.273)	A (0.494)	A (0.131)	A (0.046)
	H (0.135)	H (0.562)	H (0.087)	H (0.607)	H (0.324)
	VH (0.000)	VH (0.000)	VH (0.006)	VH (0.170)	VH (0.010)
Management capability	VL (0.000)	VL (0.000)	VL (0.000)	VL (0.000)	VL (0.000)
	L (0.047)	L (0.000)	L (0.070)	L (0.164)	L (0.036)
	A (0.317)	A (0.465)	A (0.238)	A (0.652)	A (0.667)
	H (0.565)	H (0.496)	H (0.556)	H (0.028)	H (0.097)
	VH (0.071)	VH (0.039)	VH (0.136)	VH (0.126)	VH (0.000)
Health and safety records	VP (0.000)	VP (0.000)	VP (0.000)	VP (0.112)	VP (0.000)
	P (0.244)	P (0.312)	P (0.042)	P (0.410)	P (0.000)
	I (0.510)	I (0.266)	I (0.345)	I (0.355)	I (0.200)
	G (0.206)	H (0.341)	H (0.259)	H (0.103)	H (0.250)
	VG (0.000)	VG (0.051)	VG (0.034)	VG (0.000)	VG (0.520)
Reputation	VB (0.000)	VB (0.000)	VB (0.000)	VB (0.000)	VB (0.013)
	B (0.340)	B (0.000)	B (0.065)	B (0.306)	B (0.295)
	A (0.353)	A (0.518)	A (0.459)	A (0.271)	A (0.616)
	H (0.274)	H (0.449)	H (0.387)	H (0.023)	H (0.036)
	VH (0.014)	VH (0.033)	VH (0.050)	VH (0.000)	VH (0.000)

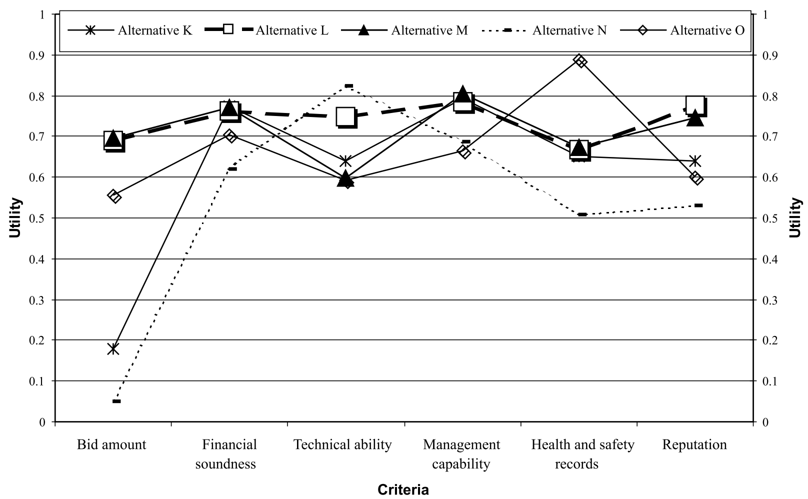

Figure 7 Aggregated assessments of contractors on six main decision criteria.

Table 1. For example, the ranking for 'bid amount' is as follows: M > L > O > K > N ($a > b$ means a is preferred to b). Note that the criterion 'bid amount' is negatively oriented in terms of preference, that is, the lower the bid amount, the better. The results in Table 1 are also useful in that they indicate the weak and strong points of each contractor regarding the decision criteria applied. The IDS provides a graphical display of the results presented in Table 1, which is very useful for DMs to compare contractors 'at a glance' (refer Fig. 7).

The assessments in Table 1 need to be propagated to the top level. In doing this, the IDS produced the results shown in Table 2. The numbers under each grade indicate the aggregated assessments (or degrees of belief) of the DM. For instance, the results for contractor K can be interpreted as follows: contractor K is assessed to be 39% *worst*, almost 30% *poor*, 13% *average*, 12% *good* and only 1% *excellent*. The total DoB does not add up to one (or 100%) as a result of incomplete and/or missing assessments. The results in Table 2 are supported by a graphical display (Fig. 8).

The contractors could be ranked in order of preference by comparing them with each other as in Table 2. However, a comparison may not be possible when contractors have very similar degrees of belief assigned to each grade, such as contractors L and M (see Fig. 8). One way to solve this problem is to quantify the grades. There are several ways of quantifying grades. One of them is to use MAUT to assign a utility for each grade and then obtain an expected utility for each contractor (cf. Hwang & Yoon, 1981). Then, contractors are ranked based on their expected utility. Another way is to use a goal programming technique such as suggested in Yang & Sen (1997). In this study, the former approach is used. A number of hypothetical lottery type questions were presented to the DM in order to establish preference among grades (Farquhar, 1984). The following utilities are assigned to each grade: worst = 0, poor = 0.4, average = 0.7, good = 0.85 and excellent = 1.

Table 2 The overall assessment of alternative contractors.

	Grades						
Alternative	Worst	Poor	Average	Good	Excellent	Total DoB*	Unassigned DoB
K	0.3908	0.2987	0.1290	0.1239	0.0105	0.9529†	0.0471‡
L	0.0000	0.0304	0.8164	0.1234	0.0094	0.9796	0.0204
M	0.0000	0.0478	0.7836	0.1220	0.0147	0.9681	0.0319
N	0.6260	0.0705	0.1716	0.0506	0.0182	0.9369	0.0631
O	0.0062	0.3688	0.4838	0.0660	0.0315	0.9563	0.0437

* Degree of belief.
† 0.3908 + 0.2987 + 0.1290 + 0.1239 + 0.0105.
‡ 1 − total DoB = 1 − 0.9529 = 0.0471.

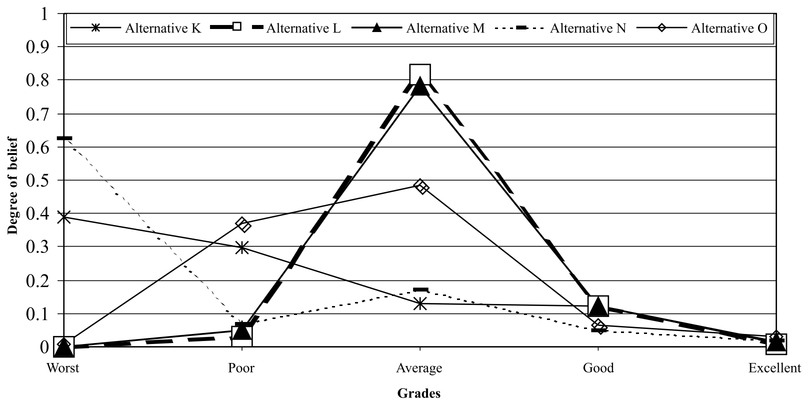

Figure 8 Graphical performance profiles of contractors.

Table 3 The expected utilities of alternative contractors.

Alternative	Utility		
	Minimum	Maximum	Average
K	0.3257	0.3727	0.3492
L	0.6981	0.7183	0.7082
M	0.6860	0.7180	0.7020
N	0.2096	0.2726	0.2411
O	0.5737	0.6175	0.5956

The total DoB for each contractor in Table 2 does not add up to one, because some of the assessments were incomplete and missing because of the reasons explained earlier in the paper. For example, the total DoB assigned to contractor M is almost 97%. That is, there is a 3% unassigned DoB. The IDS uses the concept of utility interval to characterize the unassigned DoB (or ignorance) which can actually fall into any grade. The ER algorithm generates a utility interval enclosed by two extreme cases where the unassigned DoB goes either to the least preferred grade (minimum utility) or goes to the most preferred grade (maximum utility). The minimum and maximum possible utilities of each alternative generated by the IDS (based on the given utility values for each grade above) are shown in Table 3. For example, the results for contractor M from Table 2 are as follows:

Contractor M's minimum utility = {[(DoB assigned under grade worst + unassigned DoB) × utility of grade worst] + (DoB assigned under grade poor × utility of grade poor) + (DoB assigned under grade average × utility of grade average) + (DoB assigned under grade good × utility of grade good) + (DoB assigned under grade excellent × utility of grade excellent)}. Hence, contractor M's minimum utility = $\{[(0 + 0.0319) \times 0] + (0.0478 \times 0.4) + (0.7836 \times 0.7) + (0.122 \times 0.85) + (0.0147 \times 1)\} = 0.686$.

Contractor M's maximum utility = {(DoB assigned under grade worst × utility of grade worst) + (DoB assigned under grade poor × utility of grade poor) + (DoB assigned under grade average × utility of grade average) + (DoB assigned under grade good × utility of grade good) + [(DoB assigned under grade excellent + unassigned DoB) × utility of grade excellent]}. Hence, contractor M's maximum utility $= \{(0 \times 0) + (0.0478 \times 0.4) + (0.7836 \times 0.7) + (0.122 \times 0.85) + [(0.0147 + 0.0319) \times 1] = 0.718$. Contractor M's average utility = (maximum utility + minimum utility)/2, i.e. $(0.686 + 0.718) = 0.702$.

The contractors may be ranked based on the average utility but this may be misleading. In order to say that one contractor theoretically dominates another, the preferred contractor's minimum utility must be equal or greater than the dominated contractor's maximum utility. For example, based on an average utility, contractor L is preferred to contractor M. On the other hand, this comparison may differ if it is based on the maximum and minimum utilities. There is a small possibility that contractor M may be preferred to contractor L because M's maximum utility is greater than L's minimum utility (i.e. $0.7180 > 0.6981$). To precisely differentiate between contractors L and M, the quality of the original assessments related to L and M needs to be improved. In response to the DM's request for simplicity, average utilities are used to rank contractors. The ranking of contractors is as follows: L > M > O > K > N.

CONCLUSION

In this paper, an ER approach to solve the CSP has been described that is capable of accommodating both quantitative and qualitative data. A decision-maker (i.e. a client or their representative) may be willing or able to provide only incomplete, imprecise and vague information because of time pressure, a lack of data or shortcomings in expertise when evaluating contractors against a pre-determined set of criteria. In addition, the DM may wish to evaluate intangible criteria by using linguistic variables, which facilitate the processing of raw (normally difficult to represent) data. Thus there are two problems to address: (1) how to reconcile quantitative and qualitative decision criteria (data) and (2) how to deal with incomplete information in a rational way. It is shown that the ER approach is able to tackle these two problems and can help DMs reach a robust decision although some data may be missing and/or assessments may be incomplete. A further advantage of the method is that uncertainty and risk surrounding the decision problem can be represented through the concept of 'the DoB'. The computer software IDS facilitates the implementation of the ER approach. One of the disadvantages of the method may be that it requires more complicated calculations than some other methods such as MAUT.

Selecting the best contractor is one of the most important decisions a client has to make. Conversely, it is equally important for contractors to know why their bids are rejected. In the case of a public client, the results and reasons for awarding a contractor and/or rejecting others should be explicit because of public accountability. If the client is a private one, such results might be sold to those contractors who may wish to know the reasons for their failure. In either case, the feedback of contractors' weaknesses can only help improve firms to the betterment of the industry. The IDS software based on the ER approach enables users to provide results of evaluation both in tabular and graphical forms; showing the contractors' strongest and weakest areas. Such a computer support system may also be useful, because in that a large number of contractors' data can be stored and recorded for the future use. Contractors who fail to make satisfactory bids over a period of time can be monitored and removed from the database if desired.

ACKNOWLEDGEMENTS

The first author is extremely grateful to the Ministry of Education of Turkey for the financial support. The valuable comments given by Dr N. Proudlove of Manchester School of Management are also greatly appreciated. The acknowledgement is also due to the anonymous referees of this paper for the valuable comments received.

REFERENCES

Adelman, L. (1992) *Evaluating Decision Support and Expert Systems.* John Wiley & Sons Ltd, New York.

Barron, F.H. & Barrett, B.E. (1996) Decision quality ranked attribute weights. *Management Science*, **42,** 1515–1523.

Belton, V. (1986) A comparison of the analytic hierarchy process and a simple multi-attribute value function. *European Journal of Operational Research*, **26,** 7–21.

Belton, V. & Gear, T. (1983) On a shortcoming of Saaty's method of analytic hierarchies. *OMEGA – International Journal of Management Science*, **11,** 228–230.

Belton, V. & Gear, T. (1985) The legitimacy of rank reversal – a comment. *OMEGA – International Journal of Management Science*, **13,** 143–144.

Beynon, M., Curry, B. & Morgan, P. (2000) The Dempster–Shafer theory of evidence: an alternative approach to multicriteria decision modelling. *OMEGA – International Journal of Management Science*, **28,** 37–50.

Department of Environment (1992) CMIS, *Contractor Management Information System*, The UK Department of the Environment, Transportation and the Regions, http://www.detr.gov.uk

Farquhar, P.H. (1984) Utility assessment methods. *Management Science*, **30,** 1283–1300.

Hatush, Z. & Skitmore, M. (1997a) Criteria for contractor selection. *Construction Management and Economics*, **15,** 19–38.

Hatush, Z. & Skitmore, M. (1997b) Assessment and evaluation of contractor data against client goals using PERT approach. *Construction Management and Economics*, **15,** 327–340.

Hatush, Z. & Skitmore, M. (1998) Contractor selection using multicriteria utility theory: an additive model. *Building and Environment*, **33,** 105–115.

Hipel, K.W., Radford, K.J. & Fang, L. (1993) Multiple participant–multiple criteria decision making. *IEEE Transactions on Systems, Man, and Cybernetics*, **23,** 1184–1189.

Holt, G.D. (1998) Which contractor selection methodology? *International Journal of Project Management*, **16,** 153–164.

Holt, G.D., Olomolaiye, P.O. & Harris, F.C. (1993) A conceptual alternative to current tendering practice. *Building Research and Information*, **21,** 167–172.

Holt, G.D., Olomolaiye, P.O. & Harris, F.C. (1994a) Factors influencing UK construction clients' choice of contractor. *Building and Environment*, **29,** 241–248.

Holt, G.D., Olomolaiye, P.O. & Harris, F.C. (1994b) Applying multi-attribute analysis to contractor selection decisions. *European Journal of Purchasing and Supply Management*, **1,** 139–148.

Holt, G.D., Olomolaiye, P.O. & Harris, F.C. (1994c) Incorporating project specific criteria and client utility into the evaluation of construction tenderers. *Building Research and Information*, **22,** 214–221.

Holt, G.D., Olomolaiye, P.O. & Harris, F.C. (1994d) Evaluating prequalification criteria in contractor selection. *Building and Environment*, **29,** 437–448.

Holt, G.D., Olomolaiye, P.O. & Harris, F.C. (1995a) A review of contractor selection practice in the UK construction industry. *Building and Environment*, **30,** 553–561.

Holt, G.D., Olomolaiye, P.O. & Harris, F.C. (1995b) Application of an alternative contractor selection model. *Building Research and Information*, **23,** 255–264.

Holt, G.D., Olomolaiye, P.O. & Harris, F.C. (1996) Applying cluster analysis to construction contractor classification. *Building and Environment*, **31,** 557–568.

Hwang, C.L. & Yoon, K. (1981) *Multiple Attribute Decision Making Methods and Applications: A State-of-the-Art Survey*. Springer-Verlag, New York.

Islei, G. & Lockett, A.G. (1988) Judgemental modelling based on geometric least square. *European Journal of Operational Research*, **36,** 27–35.

Keeney, R.L. & Raiffa, H. (1993) *Decisions with Multiple Objectives*. Cambridge University Press, Cambridge.

Kerry, R. (1990) *Integrated Knowledge-Based and Database Management Systems*. Ellis, Horwoood Ltd, New York.

Ng, S.T. (1993) *Three-level Intelligent Decision Model for Contractor Prequalification, Proceedings: International Symposium on Economic Evaluation and the Built Environment, CIB W55 and W95*. Laboratorio Nacional de Engenharia Civil, Lisboa, Portugal, vol. 1, pp. 119–129.

Ng, S.T. (1996) *Case-based reasoning decision support for contractor prequalification*. PhD Thesis, University of Manchester of Institute of Science and Technology, Manchester, England.

Nguyen, V.U. (1985) Tender evaluation by fuzzy sets. *Journal of Construction Engineering and Management*, **111,** 231–243.

Poyhonen, M.A., Hamalainen, R.P. & Salo, A.A. (1997) An experiment on the numerical modelling of verbal ratio statements. *Journal of Multi-Criteria Decision Analysis*, **6,** 1–10.

Russell, J.S. & Skibniewski, M.J. (1988) Decision criteria in contractor prequalification. *Journal of Management in Engineering*, **4,** 148–164.

Russell, J.S., Skibniewski, M.J. & Cozier, D.R. (1990) Qualifier-2: knowledge-based system for contractor prequalification. *Journal of Construction Engineering and Management*, **116,** 157–171.

Saaty, T.L. & Wind, Y. (1980) Marketing applications of the analytic hierarchy process. *Management Science*, **26,** 641–658.

Sen, P. & Yang, J.B. (1998) *Multiple Criteria Decision Support in Engineering Design*. Springer-Verlag, London.

Stewart, T.J. (1992) A critical survey on the status of multiple criteria decision making theory and practice. *OMEGA – International Journal of Management Science*, **20,** 569–586.

Taha, M.A., Park, S.C. & Russell, J.S. (1995) Knowledge-based DSS for construction contractor pre-screening. *European Journal of Operational Research*, **84,** 35–46.

Yang, J.B. (2001) Rule and utility based evidential reasoning approach for multiattribute decision analysis under uncertainties. *European Journal of Operational Research*, **131,** 31–61.

Yang, J.B. & Sen, P. (1994) A general multi-level evaluation process for hybrid MADM with uncertainty. *IEEE Transactions on Systems, Man, and Cybernetics*, **34,** 1458–1473.

Yang, J.B. & Sen, P. (1997) Multiple attribute design evaluation of complex engineering products using the evidential reasoning approach. *Journal of Engineering Design*, **8,** 211–230.

Yang, J.B. & Singh, M. (1994) An evidential reasoning approach for multiple attribute decision making with uncertainty. *IEEE Transactions on Systems, Man, and Cybernetics*, **24,** 1–18.

Yang, J.B. & Xu, D.L. (2000) *IDS, Version 1.1, Intelligent Decision System via Evidential Reasoning, IDSL*. Sale, Cheshire, UK.

Wang, J., Yang, J.B. & Sen, P. (1995) Safety analysis and synthesis using fuzzy sets and evidential reasoning. *Reliability Engineering and System Safety*, **47,** 103–118.

Wang, J., Yang, J.B. & Sen, P. (1996) Multi-person and multi-attribute design evaluations using evidential reasoning based on subjective safety and cost analysis. *Reliability Engineering and System Safety*, **52,** 113–128.

Zadeh, L.A. (1965) Fuzzy sets. *Information and Control*, **8,** 338–353.