European Journal of Operational Research 69 (1993) 210–220
North-Holland

Theory and Methodology

A multicriteria evaluation of methods for obtaining weights from ratio-scale matrices *

B. Golany

Faculty of Industrial Engineering and Management, The Technion – Israel Institute of Technology, Haifa, 32000 Israel

M. Kress

CEMA, P.O. Box 2250, Haifa, 31201 Israel

Received April 1991; revised July 1991

Abstract: Ratio-scale matrices are commonly used as a vehicle to distinguish among competing alternatives or conflicting criteria. The issue of extracting weights out of such matrices has attracted increasing attention in recent years and resulted in the development of several scaling methods. This paper presents an attempt to evaluate some of the more known techniques in this area. The evaluation procedure includes random generation of a large number of representative matrices which are used to evaluate the various scaling methods according to the specified criteria. The performance of the different techniques is analyzed with respect to the proposed criteria and conclusions are drawn.

Keywords: Decision theory; Ratio-scale matrices; Pairwise comparison

1. Introduction

A ratio-scale matrix $D=(d_{ij})$ is a positive square matrix with the property that $d_{ij}=1/d_{ji}$. Such a matrix is referred to also as a positive reciprocal matrix. Pairwise preference intensities between objects may be expressed in the form of such a ratio-scale matrix D. The d_{ij}-entry in D represents the extent to which object i is preferred to object j. Once such a matrix is given, the objective is to elicit from these ratio scales the implicit weights of the various objects.

If these preference intensities are consistent, that is if $d_{ij}=d_{ik}d_{kj}$ for all i, j and k, then these weights, which are unique up to a multiplication by a constant, are readily available by taking any column of D. If, however, the preference intensities are not consistent, then there is a need for a method or an algorithm for obtaining these weights.

* This research was partially supported by the E. & J. Bishop Fund through the Technion Vice President for Research Office.

Correspondence to: M. Kress, Center for Military Analyses, P.O. Box 2250 (28), Haifa, Israel.

Several such methods were proposed in the literature. These methods can roughly be divided into two classes: Eigenvector methods and Extremal methods. Following the lines of Wei's method [18], for ordinal data, Saaty [15] proposed the principle eigenvector of D as the desired weights vector of the objects. This vector is normalized. Cogger and Yu [8] proposed a variant of Saaty's Eigenvector (EV) method which is applied to upper triangular ratio-scale matrix. This method is called here Modified Eigenvector (MEV) method.

Extremal methods comprise of two major stages: First find a distance function $F(D_1, D_2)$ on the set of all ratio-scale matrices, and then for given inconsistent ratio-scale matrix D, solve the problem Min $F(D, (w_i/w_j))$ subject to some normalizing constraints on the weights w_i. The extremal methods that are considered in this paper are: Direct Least Squares (DLS), Weighted Least Squares (WLS), Logarithmic Least Squares (LLS) (known also as the Geometric Mean) and Logarithmic Least Absolute Values (LLAV).

The objective of this paper is to compare the various methods with respect to several quantitative criteria. The analysis is carried out by applying the various weight derivation methods to a sample of randomly generated ratio-scale matrices, and evaluating the results according to the selected criteria. The outcome of this analysis is an overall evaluation of the methods. Previous comparative analyses of this type are given in Chu et al. [7], Budescu et al. [6], Saaty and Vargas [16], Takeda et al. [17], and Zahedi [19]. These comparisons concentrate on analytic properties of the methods (some of these properties are repeated in this paper for completeness). In general, previous comparisons were limited to a small number of methods and usually involve 2—3 criteria to evaluate the relative performance of the methods. In most cases, the comparison itself was of a secondary importance as the main purpose of the paper was the presentation of a new method and the comparison served to provide legitimacy to the proposed methods. Chu et al. [7] compared two methods (EV, WLS) using two examples (7×7 matrices) and applying two criteria (minimum distance measure and ease of computation). Their conclusion was that none of the two dominates the other method. Budescu et al. [6] compared 4 methods (left & right EV, LLS and the LLS of the EV). They ran 30 simulations, each involving some 1000 matrices ranging from 4×4 to 10×10. They concluded that the solutions of all four were highly correlated. Saaty and Vargas [16] compared the EV, LLS and DLS methods and found the EV to outperform the other two. Their criteria were Saaty's measure of consistency [15], dual solutions and rank preservation. Their comparison is based on analytical developments which are exhibited through 500 matrices of sizes $n = 3, 4, \ldots, 13$ with randomly generated entries. Zahedi [19] compared 6 methods (EV, column & row LLS, mean transformation, harmonic mean, row sums) using two criteria (statistical accuracy and rank preservation). The comparison was based on 10 runs, each involving some 500 matrices of sizes $n = 2, 4, \ldots, 22$. The only substantial findings were that the column LLS and Row sums methods were inferior to the other four techniques. This conclusion contradicts previous findings by Crawford and Williams [11], who compared the EV and LLS and found that the LLS outperformed the EV. Their analysis used two criteria (Least Squares and Logarithmic Least Squares) and tested the two methods using 10 runs, each with 500 consistent matrices (of sizes 5, 7, and 10) in which the entries were perturbed by an error term. Takeda et al. [17] used 3 methods (EV, MEV, LLS), two distance criteria, and a simulation in which 6 runs, each involving 500–1000 matrices of size 5×5, were generated. Their finding was that the LLS was superior to the EV in some cases and equal in others.

The next section offers a brief description of each one of the aforementioned methods, along with some of its properties. In Section 3 we present the criteria which are used to evaluate the methods. The Results of the analysis are presented in Section 4, and summary and conclusions are given in Section 5.

2. The weighting methods

In this section we describe the methods that are compared in the analysis to follow. The Eigenvector methods are presented first.

Eigenvector (EV): Following the notion of 'high-order intensities' [9,15,18] and utilizing the Frobenius Theorem [13], Saaty [15] proposes the principal eigenvector of D as the desired weights vector. That is, the weights vector W is the solution of

$$D \cdot W = \mathrm{PE}(D) \cdot W, \quad e^{\mathrm{T}} \cdot W = 1 \tag{2.1}$$

where PE(D) denotes the principal eigenvalue of D. If D is inconsistent, then $\mathrm{PE}(D) > n$, where n is the order of D. Saaty defines a consistency index of a ratio-scale matrix D by

$$\mu(D) = \frac{\mathrm{PE}(D) - n}{n-1}. \tag{2.2}$$

Modified Eigenvectors (MEV): A variant of the eigenvector method, proposed by Cogger and Yu [8], is applied to upper triangular ratio-scale matrices. Since D is reciprocal, the entire preference intensity information is contained in the upper triangular matrix $T = \{t_{ij}\}$, where

$$t_{ij} = \begin{cases} d_{ij} & \text{if } j \geq i, \\ 0 & \text{otherwise.} \end{cases}$$

Define the diagonal matrix G with elements $g_{ii} = n - i + 1$. If D is consistent, then so is T, and

$$T \cdot W = G \cdot W \quad \text{or } G^{-1} \cdot T \cdot W = W \tag{2.3}$$

where W is any column of D. Since $G^{-1}T$ is triangular with the values $1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}$ on its main diagonal, it follows that there are also the eigenvalues of that matrix. A reasonable estimate for W in the inconsistent case is, therefore, the solution to the equation

$$(G^{-1} \cdot T - I) \cdot W = 0. \tag{2.4}$$

The solution for (2.4) is given recursively by

$$W_i = \sum_{j=i+1}^{n} \frac{d_{ij} \cdot W_j}{n-i}, \quad i = 1, \ldots, n-1, \tag{2.5}$$

and

$$\sum_{i=1}^{n} W_i = 1,$$

which is obviously very easy to compute.

Next we describe the extremal methods which are compared in our analysis.

Direct Least Squares (DLS): In the DLS method the objective is to find a consistent ratio-scale matrix which minimizes the Euclidean distance from D. That is,

$$\begin{aligned} \text{Min} \quad & \sum_{i=1}^{n} \sum_{j=1}^{n} \left(d_{ij} - \frac{W_i}{W_j} \right)^2 \\ \text{s.t.} \quad & \sum_{i=1}^{n} W_i = 1. \end{aligned} \tag{2.6}$$

The nonlinear optimization problem in (2.6) has no special tractable form and therefore is difficult to solve numerically.

Weighted Least Squares (WLS): The WLS method is similar to the DLS method in that it also minimizes the L^2 distance function. Specifically, we solve the following nonlinear optimization problem:

$$\begin{aligned} \text{Min} \quad & \sum_{i=1}^{n} \sum_{j=1}^{n} (d_{ij} \cdot W_j - W_i)^2 \\ \text{s.t} \quad & \sum_{i=1}^{n} W_i = 1. \end{aligned} \tag{2.7}$$

Differentiating the Lagrangian of (2.7) and equating it to zero, we obtain the following set of linear equations:

$$\begin{aligned} & \sum_{i \neq k}^{n} (d_{ik}^2 + n - 1) \cdot W_k \\ & \quad - \sum_{j \neq k}^{n} (d_{jk} + d_{kj}) \cdot W_j + \beta = 0, \quad k = 1, \ldots, n, \\ & \sum_{i=1}^{n} W_i = 1. \end{aligned} \tag{2.8}$$

While the DLS may have multiple solutions, it can be shown that the WLS solution is unique, and strictly positive (Blankmeyer, [4]).

Logarithmic Least Squares (LLS): The multiplicative nature of ratio-scale matrices leads naturally to distance functions of the type

$$F\left(d_{ij}, \left(\frac{W_i}{W_j} \right) \right) = \left\| \log d_{ij} - \log\left(\frac{W_i}{W_j} \right) \right\|. \tag{2.9}$$

A reasonable choice for that metric is again L^2. The LLS weights are obtained as the solution of

$$\text{Min} \sum_{i=1}^{n} \sum_{j=1}^{n} (\log d_{ij} - \log W_i + \log W_j)^2 \tag{2.10}$$

subject to the (multiplicative) normalizing constraint

$$\prod_{i=1}^{n} W_i = 1, \quad W_i > 0, \ i = 1,\ldots,n. \tag{2.11}$$

It can be shown (Crawford and Williams, [11]) that the unique solution of (2.10)–(2.11) is the geometric mean of the columns of D, that is,

$$W_i = \prod_{j=1}^{n} (d_{ij})^{1/n}, \ i = 1,\ldots,n. \tag{2.12}$$

Barzilai et al. [2] propose another approach, from an algebraic point of view, to justify the geometric mean method.

Logarithmic Least Absolute Values (LLAV): Cook and Kress [9] approach the concept of distance on ratio-scale matrices from an axiomatic point of view. A set of natural axioms which represent the usual metric properties and the concept of independence of irrelevant objects, uniquely determines (up to a scaling factor which w.l.o.g. is taken to be equal to 1) the LLAV as the desired distance. The problem of obtaining the weights is shown then to be equivalent to the following linear programming problem:

$$\text{Min} \sum_{i=1}^{n} \sum_{j=1}^{n} (P_{ij} + N_{ij}) \tag{2.13}$$

s.t.

$$x_i^+ - x_i^- - x_j^+ + x_j^- - N_{ij} + P_{ij} = \log d_{ij} \quad \forall i, j,$$

$$\sum_{i=1}^{n} (x_i^+ - x_i^-) = 0,$$

$$x_i^+, x_i^-, P_{ij}, N_{ij} \geq 0, \quad \forall i, j.$$

The weight vector W is given by

$$W_i = e^{x_i^+ - x_i^-}.$$

The LLAV method is essentially a median related method which is not biased towards extreme values of d_{ij}.

3. The criteria

As described in the Introduction, there is no consensus among the previous comparative studies as to which criteria are most appropriate for the evaluation of the various techniques for deriving weights from ratio-scale matrices. Almost all of the researchers recognized the multicriteria nature of the evaluation problem, but each selected a particular set of criteria. In most cases, at least one distance measure is used (see, e.g. Takeda et al. [17]). However, qualitative criteria (such as the ease of computation in Chu et al. [7]) also played an important role in these evaluations. The criteria selected for our analysis represent the spectrum of criteria used earlier, as well as additional criteria which were not applied in this context yet. The criteria are presented below, following the notation used in this paper.

Notation

$D = \{d_{ij}\}$ = Ratio-scale pairwise comparison matrix, $i, j = 1,\ldots,n$.
$w_j(t)$ = Weight allocated to alternative j, by method t, $t = 1,\ldots,T$.
$\bar{w}_j$ = Average weight allocated to alternative j by all the evaluated methods.
$r_j(t)$ = Rank position of alternative j as determined by method t.
$\rho_j(t)$ = Alternative j's rank position by method t after input perturbation.
I_{ij} = A violation indicator for the pair i, j (see (3.2) below).
$Y_j(t)$ = A position change indicator for alternative j evaluated by method t (see (3.7) below).

Minimum Violations (MV)

A 'violation' (Ali et al. [1]) or 'element preference reversal' (Jensen, [14]) occurs when alternative j is preferred to alternative i in the pairwise comparison but i receives a larger weight in the final weights vector. The MV measure sums up all the violations associated with the weight vector w for any given solution technique. That is:

$$\text{MV}_t = \sum_i \sum_j I_{ij}(t), \quad t = 1,\ldots,T \tag{3.1}$$

where

$$I_{ij}(t) = \begin{cases} 1 & \text{if } w_i(t) > w_j(t) \text{ and } d_{ij} < 1, \\ \frac{1}{2} & \text{if } w_i(t) = w_j(t) \text{ and } d_{ij} \neq 1 \\ & \text{or } w_i(t) \neq w_j(t) \text{ and } d_{ij} = 1, \\ 0, & \text{otherwise.} \end{cases}$$

An important property of any weighting method is the ability to preserve the ordinal preferences which are implicitly expressed by the ratio-scale preference matrix entries. This criterion relates the ordinal deviation between the derived weights and the ratio-scale data, therefore supplying a measure of ordinal consistency. In the context of ordinal preference structures (tournaments) the axiomatic foundation of the MV as a criterion or measure for selecting the corresponding ordinal ranking is given in Ali et al. [1].

Total Deviation (TD)

The TD is a quadratic error measure which evaluates the sum of the deviations between ratios of weights and their corresponding entry in the matrix D (see Takeda et al. [17]). It is computed as

$$\mathrm{TD}_t = \sum_i \sum_j \left(d_{ij} - \frac{w_i(t)}{w_j(t)} \right)^2, \quad t = 1, \ldots, T. \tag{3.3}$$

The L^2 distance is a reasonable and widely used measure of deviations. It is natural to expect that the DLS (which in fact employs the TD measure) and the WLS methods will rate best according to this criterion. The TD criterion measures the actual (euclidean) distance between the ratios obtained by the derived weights, and the ratio-scale raw data.

Conformity (C)

The deviation between the weights vector $W(t)$ obtained by each method t and an average vector $\bar{w}$ across all the methods, measures the conformity of this method with the rest. Arguably, the larger this measure is for a particular method, the more it can be considered as an 'outlier' with respect to the other methods. In addition to identifying outliers this criterion may indicate clusters of methods which behave similarly. The measure is computed as

$$C_t = \sum_j | w_j(t) - \bar{w}_j |, \quad t = 1, \ldots, T \tag{3.4}$$

where

$$\bar{w}_j = \frac{1}{T} \sum_t w_j(t). \tag{3.5}$$

Robustness (R)

This measure is used to analyze the sensitivity of the solution methods to small changes in the input data (the matrix D). The measure is defined by the number of changes in the relative positioning of the alternatives that resulted from the input perturbation (see Section 4 for the details of generating the changes in the input matrix). Define

$$R_t = \sum_j Y_j(t), \quad t = 1, \ldots, T \tag{3.6}$$

where

$$Y_j(t) = \begin{cases} 1 & \text{if } r_j(t) \neq \rho_j(t), \\ 0 & \text{otherwise.} \end{cases} \tag{3.7}$$

In addition to the four aforementioned quantitative criteria it is also possible to compare the methods by a variety secondary, more qualitative, criteria. Some of these are:

Alternative Solutions (AS): Some methods may produce multiple solution vectors. This feature may allow for sequential optimization (or 'prioritized goal programming'). The alternate weight vectors can be searched (as a secondary analysis stage) to find the one which best satisfies a given criterion. For example, as it is shown in the next section, the LLAV method best satisfies the Minimum Violation criterion. It may also produce alternative solutions. Out of the several possible MV solutions one can select the weight vector which is best with respect to the TD criterion.

Computation Time (CT): The number of CPU seconds needed to execute the solution techniques can be viewed as one of the performance criteria. Obviously, methods which require the execution of optimization routines (e.g., DLS, LLAV and WLS) are inferior to methods which can be executed via simple arithmetic calculations (e.g., LLS).

Invariance to Transposition (IT): A desirable property of any solution method to the investigated problem is invariance to the transposition operation on the elements of the input matrix D. In other words, if a permutation is applied to the order of the objects in D, then the result should not change except for the same permutation in the order of the objects (see Barzilai et al. [2], Axiom 2).

4. Results of the analysis

A large number of ratio-scale matrices were generated by a simulation program which was created specifically for the purpose of our comparative study. The program [1] takes about a minute to complete a cycle which includes generating a random matrix of size 7 × 7, applying the six solution techniques to the matrix, evaluating the different criteria and storing the outputs for future presentation.

The matrices were generated by a random generator procedure which worked as follows. First, the entries in the first row were randomly selected from a uniform distribution in the interval [1,9]. Thus, without loss of generality, the first row corresponds to an alternative which is preferred (or equal) to all the other alternatives:

$$d_{1j} \sim U(1, 9), \quad j = 1, \ldots, n.$$

[1] Written in Turbo Pascal and executed on an IBM-AT PC, the program uses the GAMS package as an internal module for solving the methods which require the solution of mathematical programs.

If the matrix is consistent, all the other entries in it can be computed directly from the first row. A critical issue in any such analysis is how to model the inconsistency which is inherent in human judgments. Traditionally, choices were made between additive or multiplicative error terms. The procedure here is different, although it bears some resemblance to the one used by Takeda et al. [17] and Zahedi [19]. A user-driven parameter, denoted as k (%), determines an interval around the consistent value (computed from the first row information). The parameter k is referred to as 'degree of inconsistency'. The entry in the matrix is then randomly selected from a uniform distribution in that interval. Specifically,

$$d_{ij} \sim U\left(\frac{(100-k) \cdot d_{1j}}{100 \cdot d_{1i}}, \frac{(100+k) \cdot d_{1j}}{100 \cdot d_{1i}} \right),$$

$$j > i > 1$$

and

$$d_{ji} = \frac{1}{d_{ij}}.$$

Thus, for example, if $d_{12} = 2$, $d_{13} = 4$, and $k = 90\%$, d_{23} is selected from the interval [0.2, 3.8]. The matrices that were generated for the analysis were of sizes $n = 5, 6, 7$. For every value of n, 9 values of k (10, 20, ..., 90) were applied and each combination of $\{n, k\}$ was replicated 30 times, thus resulting in a total of 810 matrices.

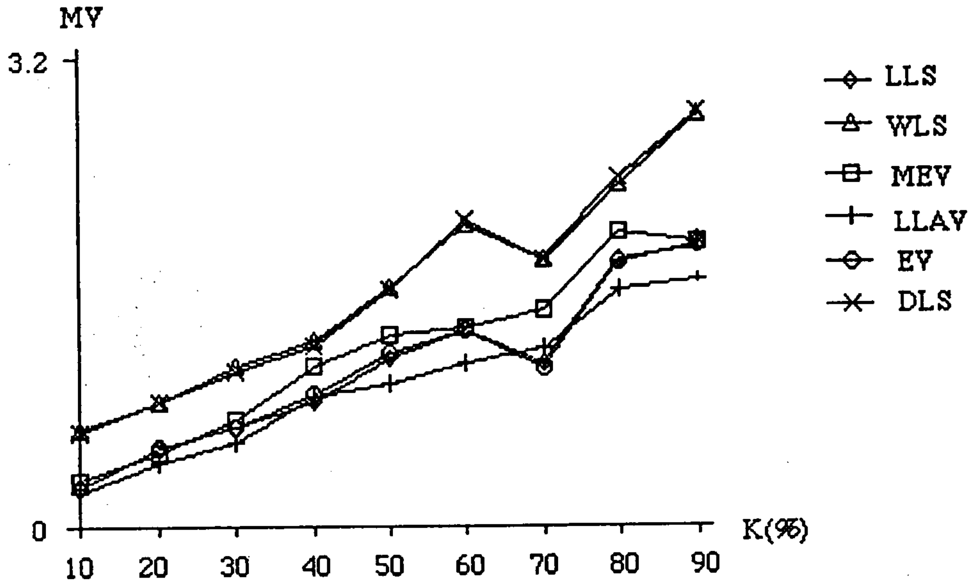

Figure 1. Minimum violations vs. degree of inconsistency

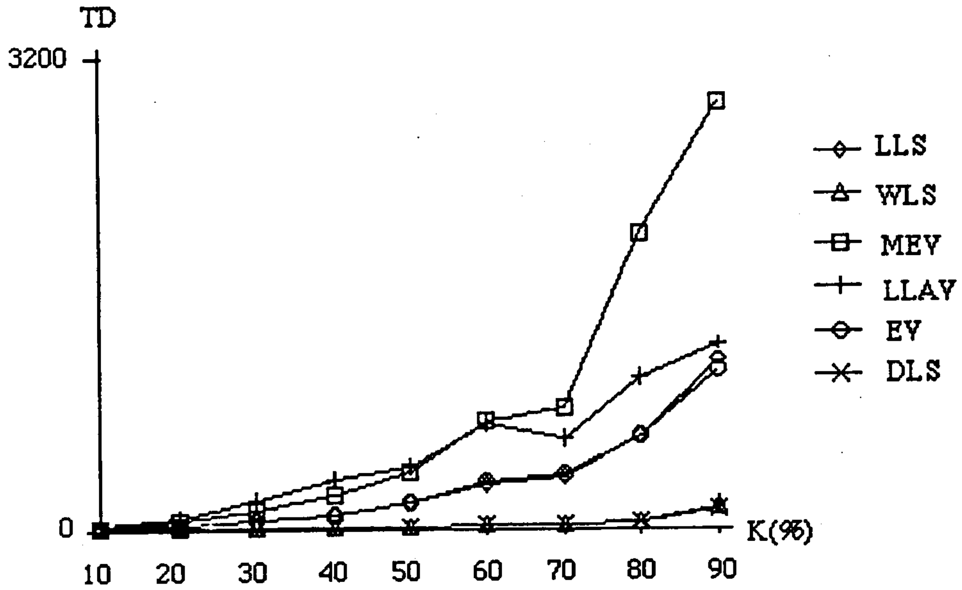

Figure 2. Total deviations vs. degree of inconsistency

The generation of matrices for the sensitivity analysis tests was slightly different. First, an initial matrix was generated and evaluated as explained above. Then, 10 matrices were generated on the basis of the initial matrix by allowing random changes in intervals of ± 10, 20, 30, 40, 50% from the initial matrix's values. The sensitivity module of the analysis was replicated three times. In each run, 9 initial matrices were generated (with k-values $10, 20, \ldots, 90$) and each initial matrix was branched into 10 related matrices as explained above. Thus, a total of 270 matrices were evaluated in that manner.

Next, we present the results of the analysis with respect to each one of the criteria given in Section 3. Each point in the graphs shown in Figures 1–5 represents an average taken over 90 matrices (with a fixed k-parameter) of an outcome associated with one weighting method.

With regard to the AS criterion, it is observed that all the methods except two provide a unique weighting vector. The LLAV and DLS methods may have multiple solutions. Since each one of the optimization procedures associated with these two methods produces an arbitrary optimal solution, it follows that the corresponding outcomes

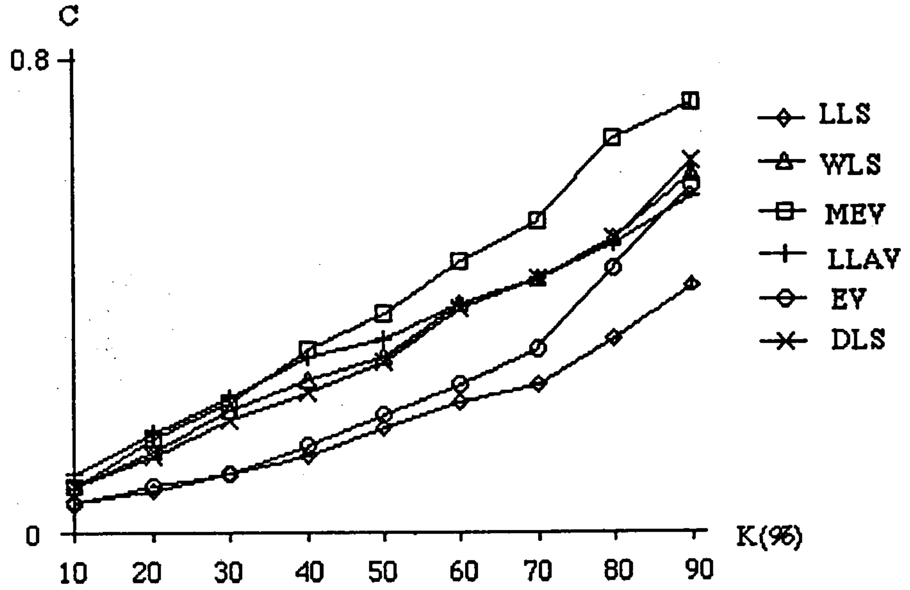

Figure 3. Conformity vs. degree of inconsistency

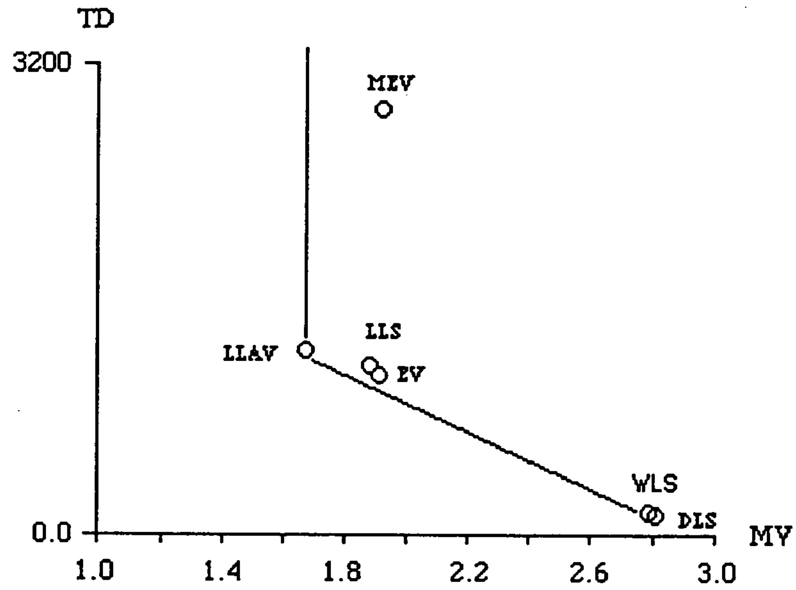

Figure 4. Minimum violations vs. total deviations (for $k = 90$)

for these methods in the graphs represent upper bounds for

a) the minimum TD- and C-values for the LLAV method, and

b) the minimum MV- and C-values for the DLS method.

Mininum violations: Figure 1 presents a graphical representation of the MV criterion. It suggests that the six methods can be categorized into three groups with respect to that criterion. In the leading group we find only LLAV. The following group includes EV, LLS and MEV whose outcomes are hardly distinguishable. The last group consists of the WLS and DLS method. It should be noted again that the DLS outcome in the graph is an upper bound to the actual MV value.

Total deviation: As for the TD criterion, the worst performer is consistently the MEV method. It is followed by EV, LLS and LLAV. The best results were obviously achieved jointly by DLS and WLS. Again, the LLAV outcomes represent upper bounds for the actual minimum TD values.

Conformity: The outcome in terms of the conformity criterion is similar. The MEV method is

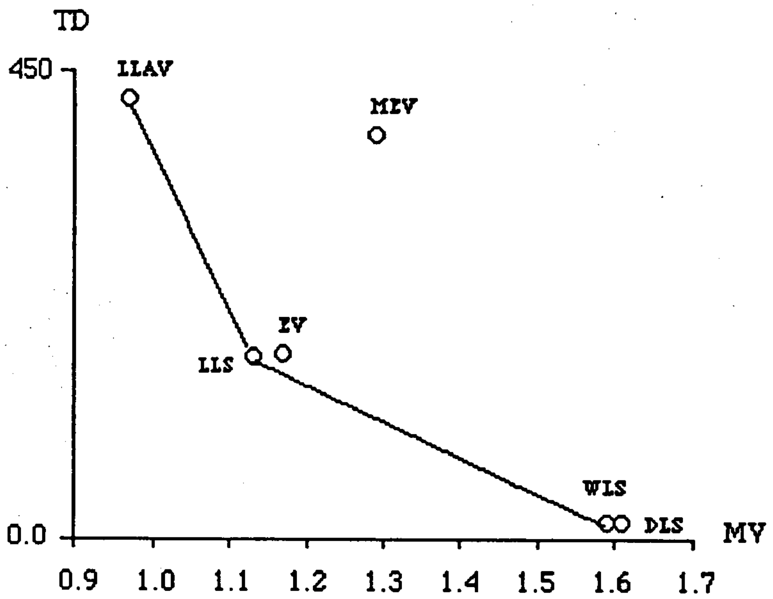

Figure 5. Minimum violations vs. total deviations (for $k = 50$)

Table 1
Sensitivity analysis in three runs

Run\Method	EV [a]	MEV [a]	LLS [a]	DLS [b]	LLAV [b]	WLS [a]
1	73	90	65	40	132	64
2	63	86	73	85	127	93
3	61	87	84	107	81	84
Total	197	263	222	232	340	241

Note that these results represent an exact performance measure for the methods designated by [a] and an 'average' performance measure for those designated by [b].

the furthest from the average (i.e., it is an 'outlier' with respect to all other methods). The group LLAV, WLS and DLS follow with almost equal results. The EV and LLS methods lead (i.e., are closest to the average), where the latter is slightly closer to the average than the former.

Figures 4 and 5 analyze the same results which were presented above one criterion at a time, in two-dimensional, 'efficiency frontier', manner. Holding k fixed, these graphs illustrate the fact that some methods outperform other methods in a particular criterion, (e.g., LLAV in MV), but are outperformed by others in terms of another criterion (again, LLAV in TD). However, their relative advantage in one or more criteria makes them 'undominated' by other methods. At the same time, other methods, (e.g., MEV), are clearly dominated by the frontier constructed by all the participating methods.

Robustness: As explained earlier, the robustness of each method was evaluated by counting the number of position changes in the ordered vector of alternatives as affected by applying each method to a series of matrices in which small perturbation from an initial matrix were implemented. Table 1 summarizes the outcome of 3 such runs, where in each run 9 initial matrices (of size 7×7) were generated and each one of them was branched (into a different perturbation) 10 times. The entries in the table correspond to the number of positional changes that were observed after the input perturbations. Since the 'worst' case is that all the alternatives change their relative position, the maximum possible entry (for each run) in Table 1 is 630 counts. The table suggests that the EV technique is the less sensitive to such changes. It is followed by the group MEV, LLS, DLS and WLS, while the LLAV technique is singled out as the most sensitive method.

Computation time: As expected the computation time turned out to be totally determined by the problems size (n). The group WLS, DLS and LLAV which call for the solution of mathematical programs required almost identical times. These were an order of magnitude larger than the CPU times needed for the next group, EV and MEV, which require an iterative procedure (implemented here within the Pascal program). The LLS technique has a clear advantage in this criterion since all it needs is a straight arithmetic manipulation which can easily be done even for very large matrices.

Alternative solutions: As for the AS criterion, it turned out that the number of the alternate solution vectors produced by LLAV ranged from two to eight for the (7×7)-matrices (similar results were obtained for the DLS method). This relatively large number can be explained by the structure of the LLAV constraints which offer the

Table 2
Summary of results

	MV	TD	*C*	*R*	CT	AS	IT
LLS	0	0	+	0	+	−	+
WLS	−	+	0	0	−	−	−
MEV	0	−	−	0	0	−	−
LLAV	+	0	0	−	−	+	+
EV	0	0	+	+	0	−	−
DLS	−	+	0	0	−	+	+

program alternative means to reach the same sum of total deviations in the objective function. As the alternative optimal weight vectors differed drastically in their values, it was established that indeed they are associated with quite different criteria values. However, due to the complex computational requirements involved in this particular secondary analysis, it was decided not to carry it within the scope of the comparison presented here.

Invariance to transposition: The LLS, DLS and LLAV methods comply with this property while the EV, MEV and WLS methods do not.

Finally, an overall look at the results is given in Table 2 below where the symbols (+), (0) and (−) correspond to good, mediocre and poor performance, respectively, for any combination of method and criterion. The table is deliberately not 'weighted' in terms of the relative importance of the different criteria since this issue is not within the scope of this work.

Conclusions

This paper describes a comprehensive study of several methods which estimate weights from ratio-scale matrices. The objective of the work was to select a group of well known methods, identify a number of relevant 'metric' measures which can serve as the instruments in comparing these methods. Six methods were selected and compared according to these four measures. In addition to the four primary criteria, three secondary criteria were also considered for evaluating the six methods' performance. Compared with other similar evaluations which have appeared so far (e.g., Zahedi [19], Budescu et al. [6], and Chu et al. [7]), this study is more comprehensive both in terms of the number of analyzed methods and the number of criteria.

Of the six compared methods, only the modified eigenvalue (MEV) method is singled out as ineffective across the different criteria. The remaining five methods have different weaknesses and advantages but none of them is dominated by the other. Special attention should be given to the criteria in which alternative solutions are possible. As stated in the text, in these cases our results provide only a bound on the methods' performance.

The choice of method for obtaining the weights should be dictated by the objectives of the analysis and the desired measure of effectiveness. Evidently, different objectives may result in different scaling methods.

Occasionally, some elements of the matrix D may be absent due to the decision maker's inability to provide a preference for every possible pair. The weighting methods may be compared with respect to their ability (or lack thereoff) to handle these situations. It should be noted that the Eigenvalue-based methods do not work in such cases, while others, can still be used. The LLS can be applied by averaging over a smaller number of elements, and the distance based methods (e.g., the DLS) can work when we assume that the missing elements do not contribute to their relevant objective functions.

Acknowledgement

The authors wish to express their gratitude to Mr. A. Kleinstein and S. Ginzborg for their help in performing the necessary computations.

References

[1] Ali, I., Cook, W.D., and Kress, M., "On the minimum violations ranking of a tournament", *Management Science* 32 (1986) 660–672.

[2] Barzilai, J., Cook, W., and Golany, B., "Consistent weights for judgements matrices of the relative importance of alternatives", *Operations Research Letters* 6 (1987) 131–134.

[3] Belton, V., and Gear, T., "On a short coming of Saaty's method of analytic hierarchies", *OMEGA* 11 (1983) 228–230.

[4] Blankmeyer, E., "Approaches to consistency adjustments", *Journal of Optimization Theory and Applications* 54 (1987) 479–488.

[5] Budescu, D.V., "Scaling binary comparison matrices: A comment on Narasimhan's proposal and other methods", *Fuzzy Sets and Systems* 14 (1984) 187–192.

[6] Budescu, D.V., Zwick, R., and Rapoport, A., "A comparison of the Eigenvalue method and the Geometric Mean Procedure for ratio scaling", *Applied Pschylogical Measurement* 10 (1986) 69–78.

[7] Chu, A.T.W., Kalaba, R.E., and Spingarn, K., "A comparison of two methods for determining the weights of belonging to fuzzy sets", *Journal of Optimization Theory and Applications* 27 (1979) 531–538.

[8] Cogger, K.O., and Yu, P.L., "Eigenweight vectors and least distance approximation for revealed preference in pairwise weight ratios", *Journal of Optimization Theory and Applications* 46 (1985) 483–491.

[9] Cook, W.D., and Kress, M., "Deriving weights from pairwise comparison ratio matrices: An axiomatic approach", *European Journal of Operational Research* 37 (1988) 355–32.

[10] Cook, W.D., and Kress, M., "An m-th generation model for weak ranking of players in a tournament, *Journal of the Operations Research Society* 41 (1991) 1111–1119.

[11] Crawford, G., and Williams, C., "A note on the analysis of subjective judgement matrices", *Journal of Mathematical Psychology* 29 (1985) 387–405.

[12] Forman, E.H., "Relative vs. absolute worth", *Mathematical Modeling* 9 (1987) 195–202.

[13] Frobenius, G., "Über matrizen aus nicht negativen elementen", *Sitz. Preuss Akad. Wiss.* (1912) 456.

[14] Jensen, R.E., "Consistency adjustment surrogacy of human response in AHP analysis: Issues, controversies, and state of the art", Working Paper 182, Trinity University, TX, 1989.

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

[16] Saaty, T.L., and Vargas, L.G., "Comparison of eigenvalues, logarithmic least squares and least squared methods in estimating ratios", *Mathematical Modeling* 5 (1984) 309–324.

[17] Takeda, E., Cogger, K.O., and Yu, P.L., "Estimating criterion weights using eigenvectors: A comparative study", *European Journal of Operational Research* 29 (1987) 360–369.

[18] Wei, T.H., "The algebraic foundation of ranking theory", Unpublished Thesis, Cambridge University, 1952.

[19] Zahedi, F., "A simulation study of estimation methods in the Analytic Hierarchy Process", *Socio-Economic Planning Sciences* 20 (1986) 347–354.