Uhandisi Journal Vol. 21 No. 2, September 1997, pp 72-84

A MEASURE OF CONSISTENCY IN THE ANALYTIC HIERARCHY PROCESS: AN EXTENSION OF GOLDEN AND WANG METHOD

Christian B. Alphonce,
University of Dar es salaam, Department of Mathematics
P. O. Box 35062, Dar es salaam, Tanzania.

ABSTRACT

One of the aspects of the Analytic Hierarchy Process which has made it a popular tool in Multicriteria Decision Making is its ability to measure the consistency in decision makers' judgements. A number of methods for doing this have been proposed by various researchers over the years. This paper extends the method of Golden and Wang which is based on additive normalisation of the priority vector. The proposed extension is based on multiplicative normalisation of the priority vector.

INTRODUCTION

The Analytic Hierarchy Process (AHP) proposed by Saaty [1] more than two decades ago has been a popular decision support tool in Multicriteria Decision Making problems. The method has been applied to diverse areas such as Systems Engineering, Operations Research, Management Science, Economic Planning, etc.

The AHP technique basically involves the following stages:

a) decomposition of the problem into a hierarchy of goals, subgoals, criteria and alternatives,
b) pairwise comparison of items at any level of the hierarchy with respect to their relative impact or contribution towards those items at the immediate higher level, and
c) using composition to estimate the relative weights of importance of items at each level.

The ability of the AHP to measure the consistency in decision makers' judgements has made it a preferred choice to many practitioners over other Multicriteria Decision Making tools. This paper begins by briefly reviewing some of the methods proposed for measuring the consistency of decision makers' judgements in the AHP. This is followed by a description of the preferable measure. The paper ends with an illustrative example on how to apply the proposed measure and recommendations for further research.

SOME METHODS OF MEASURING CONSISTENCY IN THE AHP.

The Traditional Eigenvalue Method.

Suppose **A** is an n x n pairwise comparison matrix (judgement matrix) supplied by the decision maker. The traditional eigenvalue method is based on the fact that if **A** is perfectly consistent, then the principal eigenvalue of **A** is equal to n. If **A** is not perfectly consistent then its principal eigenvalue will be greater than n.

Because the principal eigenvalue of a pairwise comparison matrix approaches n as the matrix approaches perfect consistency, Saaty [1] proposed a consistency index, C.I., defined as:

$$C.\,I. = \frac{\lambda_{max} - n}{n - 1} \qquad (1)$$

where λ_{max} is the principal eigenvalue of matrix **A**. This index is taken as a measure of consistency or reliability of judgements supplied by the decision maker.

Since for matrices of different sizes no comparisons can be made with their C.I.s', this led Saaty to define the Consistency Ratio, C.R., which is a ratio of the matrix C.I. to a Random Index, R.I. The Random Index is defined as an average of C.I.s of randomly generated reciprocal matrices with the same size. That is:

$$C.\,R. = \frac{C.\,I.}{R.\,I.} \qquad (2)$$

The Consistency Ratio is taken as a measure of the consistency of **A** which is independent of the size of the matrix. Through experience Saaty recommended that for **A** to be acceptable as being fairly consistent, its C.R. should not exceed ten percent.

A number of researchers have criticised this measure of consistency. Islei and Lockett [2] suggested that the above method provides a crude measure with limited statistical properties. Golden and Wang [3] argued that the consistency ratio and the ten percent cut - off rule are somewhat arbitrary. Barzilai [4] questions the sense in which, for two matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ of the same size, $\lambda_{max}(\mathbf{A}_1) < \lambda_{max}(\mathbf{A}_2)$ corresponds to $\mathbf{A}_2$ being more consistent than $\mathbf{A}_1$ when both matrices are inconsistent.

Other criticisms on this method include its dependence on the scale used and also its dependence on the principal eigenvector method of estimating weights of elements involved in a pairwise comparison. It has been observed [3] that this method is easily satisfied by smaller matrices but difficult to satisfy with large matrices. These criticisms point out the need to develop other more robust methods which are independent of scale and method of extracting weights and which behaves in the same way for matrices of the same size.

The Ordinal Consistency Approach.

In an effort to avoid the weaknesses of the traditional eigenvalue method of measuring consistency in the AHP, Liang and Sheng [5] proposed a procedure of detecting consistency in a pairwise comparison matrix based on the concept of ordinal consistency.

Liang and Sheng claim that a pairwise comparison matrix **A** should only be used to estimate underlying weights of elements involved in the comparisons if the binary relation used to perform the comparisons is transitive.

Since ordinal consistency of **A** implies transitivity of the used binary relation, it suffices to establish the ordinal consistency of **A** as measure of its acceptability as being fairly consistent. The advantage of this method over others is its ability to point to entries in the pairwise comparison matrix, responsible for inconsistency on top of rejecting the matrix as inconsistent. In this way a decision maker can only be asked to review his/her judgements resulting in particular entries.

The major limitation of this method is that when it comes to estimating the consistency of the entire hierarchy there is no means of doing it.

The Absolute Deviations Approach.

This method was proposed by Golden and Wang [3] in an effort to find a method free from the limitations of the traditional eigenvalue approach. This method is based on the fact that suppose **A** is a pairwise comparison matrix supplied by the decision maker and **g** is a vector of its row geometric means. If vector **g** is additively normalised a new vector **g*** results. Suppose also each column of **A** is additively normalised to give a new column $\mathbf{A}_j^*$ Golden and Wang [3] proposed the use of the mean of absolute deviations of $\mathbf{A}_j^*$ from **g*** as a measure of consistency of **A**.

It is obvious that if **A** is perfectly consistent then the mean of absolute deviations of $\mathbf{A}_j^*$ from **g*** equals zero. The closer the mean of absolute deviations of $\mathbf{A}_j^*$ from **g*** is to zero indicates the closer **A** is to perfect consistency.

To establish the cut - off value for this measure, they conducted a simulation experiment in which 1000 intelligent decision makers were each filling in a n x n pairwise comparison matrix. For each decision maker the value of the mean of absolute deviations of $\mathbf{A}_j^*$ from **g*** was computed and its frequency distribution studied. This was repeated for $n = 3$ to $n = 11$. They found that the frequency distribution of the mean of absolute deviations of $\mathbf{A}_j^*$ from **g*** was approximately Normal for $4 \leq n$. From their computational experiences they recommended the 33rd percentile of the distribution of the mean of absolute deviations of $\mathbf{A}_j^*$ from **g*** as a cut-off value for each value of n.

The advantage of this method is that it behaves almost in the same way for matrices of all sizes. It is not dependent on the method of extracting underlying weights used. Its use of additive normalisation of the vector **g** and the columns of the matrix **A** will make a number of people uncomfortable with it, especially those who attribute most of AHPís shortcomings, such as rank reversal and inverse inconsistency to its use of additive normalisation.[6]

To resolve this, this study proposes a variant of the Golden and Wang approach which uses multiplicative normalisation. Because it have been demonstrated that the AHP problem basically has a multiplicative structure [7]

THE PROPOSED MEASURE OF CONSISTENCY.

Suppose the pairwise comparison matrix $\mathbf{A}=[a_{ij}]$ supplied by the decision maker is perfectly consistent, that is each $a_{ij}=w_i/w_j$, where w_i is the weight (or relative importance) of element i involved in the pairwise comparisons for i = 1, 2, ..., n. The matrix **A** can then be represented as follows:

$$\mathbf{A}=\begin{bmatrix} w_1/w_1 & w_1/w_2 & \cdots & w_1/w_n \\ w_2/w_1 & w_2/w_2 & \cdots & w_2/w_n \\ \vdots & & & \\ w_n/w_1 & w_n/w_2 & \cdots & w_n/w_n \end{bmatrix} \tag{3}$$

The geometric means approach [6] with multiplicative normalisation yields the foll(

$$\mathbf{g} = [g_i] \quad \text{where } g_i = \frac{w_i}{\left[\prod_{j}^{n} w_j\right]^{\frac{1}{n}}} \quad \text{and} \quad \prod_{i=1}^{n} g_i = 1 \tag{4}$$

Now suppose each of the columns of **A** is normalised multiplicatively. This will give a new column normalised matrix $\mathbf{A}^*$ which is represented as follows:

$$\mathbf{A}^*=\begin{bmatrix} w_1/D & w_1/D & \cdots & w_1/D \\ w_2/D_1 & w_2/D & \cdots & w_2/D \\ \vdots & & & \\ w_n/D & w_n/D & \cdots & w_n/D \end{bmatrix} \tag{5}$$

Where

$$D = \left[\prod_{i=1}^{n} w_i\right]^{\frac{1}{n}} \tag{6}$$

If we define

$$d = \frac{1}{4n^2}\sum_i \sum_j \left|a^*_{ij} - g_i\right| \tag{7}$$

Then from equations (8), (9) and (10) it is obvious that d = 0 when **A** is perfectly consistent. As **A** becomes less consistent the value of d increases from zero. This means d can be used as a measure of consistency of judgement matrices supplied by decision makersí in the AHP.

The choice of $1/4n^2$ as a multiplying factor in (7) was arrived at after experimenting with a variety of factors. It turned out that the probability distribution of d when $1/4n^2$ is used as a multiplying factor approximated better to a normal distribution than with other experimented possibilities as will be explained later.

To be able to use d as described above there is a need to establish a cut - off value beyond which a judgement matrix will not be acceptable as being fairly consistent. To achieve this a simulation experiment similar to the one done by Golden and Wang [3] was performed. In this experiment it was assumed that 1500 knowledgeable decision makers were filling a n x n judgement matrix as follows:

1) Each a_{ii} entry, for i = 1, 2, ..., n, is assigned a value of 1.
2) Set **H** is made empty.
3) The value of each a_{1j} entry, for j = 2, 3, ..., n, is randomly picked from the set:,

$$\mathbf{S} = \left\{\frac{1}{9}, \frac{1}{8}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, 1, 2, 3, 4, 5, 6, 7, 8, 9\right\}$$

and $a_{j1}=1/a_{1j}$.

4) For i = 2, 3, ..., n-1 and j = i+1, ...,n the following is done:
   a) the ratio a_{1j}/a_{1i} is computed,
   b) Three small elements greater than the ratio are selected from **S** and stored in set **H**.
   c) Three large elements less than the ratio are also selected from

S and added to **H.**

The rationality of picking three elements from either side of the computed ratio is based, as explained by Golden and Wang [3], on the observation that, the probability distribution we are trying to construct requires the following: Suppose the decision maker is trying to be consistent will he succeed? Picking three elements indicates that the decision maker is really trying to be consistent.

Choosing 4 or 5 elements from either side of the computed ratio is okay for less sophisticated decisions. Choosing 1 or 2 elements from either side of the computed ratio will force the decision maker to be only consistent.

d) If the ratio is an element of **S** it is included in **H.**
e) If we do not get three elements in b (or c) then add the missing elements to the number of elements to be selected in c (or b).
f) Finally a_{ij} is assigned a value randomly picked from **H** and $a_{ji} = 1/a_{ij}$

5) Then d is computed using equation (11).

The above experiment was repeated for n = 3 to n = 10. This is because matrices of size greater than 10 are rare in real life AHP applications. When they occur clustering can be employed to reduce their sizes. The frequency histograms for the distribution of d are shown in Figure 1. In this Figure the observed values of d for each n are along the horizontal axis and their respective frequencies are along the vertical axis.

The above simulation experiment was performed using a computer program written in Turbo Pascal for DOS Version 6.0. This program is available from the author on request.

From the histograms in Figure 1 it looks as if the distributions of d are Normal. Goodness of fit tests for these distributions of d, assuming they were Normal, was conducted for each n. The results of these tests are summarised in Table 1:

From Table 1 we see that d is normally distributed for $4 \leq n$. The 33rd percentile of the Normal distribution of d is taken as the cut - off value. See Golden and Wang [3] for the justification of

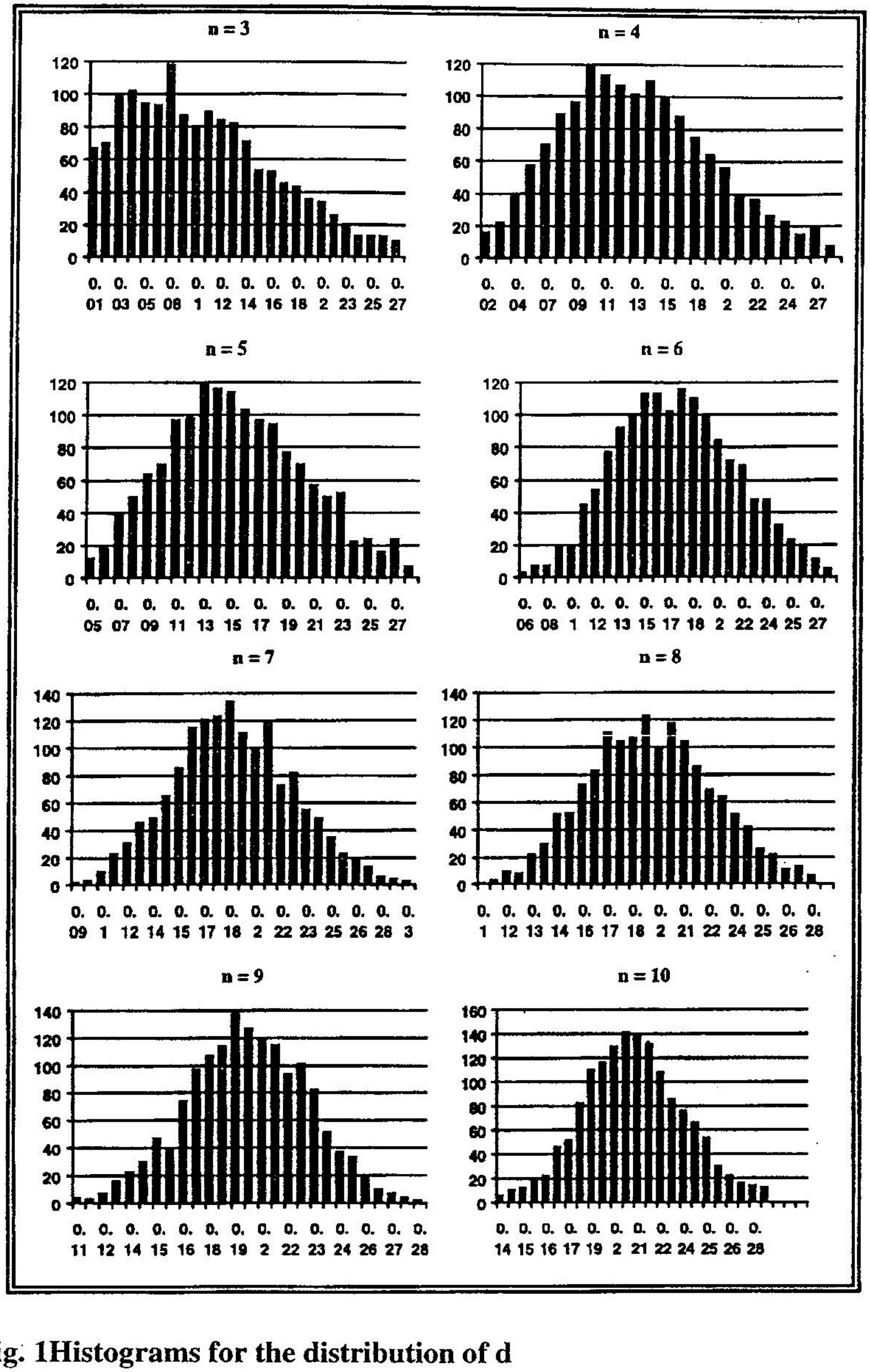

Fig. 1Histograms for the distribution of d

Table 1: Normality Tests for the Distributions of d.

n	Computed χ^2 - value	Degrees of Freedom	Critical χ^2 - value (at $\alpha = 0.05$)
3	85.7417	23	35.172
4	30.3788	22	33.924
5	27.5239	22	33.924
6	33.8845	24	36.415
7	30.7016	24	36.415
8	31.7914	25	37.652
9	21.5285	24	36.415
10	24.2831	21	32.671

this choice.

This means, a pairwise comparison matrix A= $[a_{ij}]$ supplied by a decision maker is accepted as being fairly consistent if its d value does not exceed the 33rd percentile of the Normal distribution of d for matrices of its size. Otherwise the matrix is rejected and judgement revisions are sought. Table 2 gives the cut - off values for matrices of sizes from 3 to 10 as computed from the simulation experiment.

Table 2: Cut - off Values for the Proposed Measure of Consistency

n	Average d value	Standard Deviation	33rd Percentile
3	0.09291	0.06772	0.060570
4	0.12644	0.05826	0.096700
5	0.14816	0.05139	0.122878
6	0.16903	0.04466	0.146814
7	0.18151	0.03824	0.163298
8	0.19011	0.03406	0.173500
9	0.19914	0.03406	0.185009
10	0.20604	0.02792	0.192930

EXAMPLE OF USING THE PROPOSED MEASURE OF CONSISTENCY.

Saaty's School Selection Example [1] is used to illustrate the use of the proposed measure of consistency. In this example the decision maker wanted to compare three high schools A, B and C according to their desirability.

Six criteria were used in the comparisons, they were: learning, friends, school life, vocational training, college preparation and music classes, abbreviated as L, F, S, V, C and M respectively. Table 3 give the pairwise comparison matrix obtained from comparing the criteria with respect to overall satisfaction with school.

Table 3: Pairwise Comparison Matrix for the Relative Importance of Criteria.

	L	F	S	V	C	M
L	1	4	3	1	3	4
F	$\frac{1}{4}$	1	7	3	$\frac{1}{5}$	1
S	$\frac{1}{3}$	$\frac{1}{7}$	1	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{6}$
V	1	$\frac{1}{3}$	5	1	1	$\frac{1}{3}$
C	$\frac{1}{3}$	5	5	1	1	3
M	$\frac{1}{4}$	1	6	3	$\frac{1}{3}$	1

The multiplicatively normalised row geometric means solution for the matrix in Table 3 yields the following priority vector:

g = (2.289428, 1.008165, 0.261185, 0.906681, 1.709976, 1.069913)

for the criteria L, F, S, V, C and M respectively. Normalising multiplicatively each of the columns of the pairwise comparison matrix in Table 3 leads to the matrix in Table 4.

Table 4: Column Multiplicatively Normalised Pairwise Comparisons Matrix.

2.28943	4.03266	0.78355	0.90668	5.12993	4.27965
0.57236	1.00816	1.82829	2.72004	0.34199	1.06991
0.76314	0.14402	0.26118	0.18134	0.34199	0.17832
2.28943	0.33605	1.30593	0.90668	1.70997	0.35664
0.76314	5.04082	1.30593	0.90668	1.70997	3.20974
0.57237	1.00816	1.56711	2.72004	0.56999	1.06991

From **g** and the column multiplicatively normalised pairwise comparisons matrix in Table 4 we get, for the supplied judgement matrix, the value of d is 0.193877. This value is greater than the cut - off value for a 6 x 6 judgement matrix given in Table 2.

This indicates that the supplied judgement matrix can not be accepted as being fairly consistent. This is quite in agreement with the eigenvector method which gives the C.R. of the given judgement matrix as 0.24 well above the ten percent cut - off value.

The computation of the hierarchy consistency using this proposed measure is the same as that proposed by Golden and Wang [3]. That is, first compute the hierarchy d value as:

$$\textbf{Hierarchy d value} = \frac{\sum_{i=1}^{h}\left(d_i * g_i^*\right)}{\sum_{i=1}^{h} g_i^*} \tag{8}$$

where d_i is the d value of matrix i, g_i^* is the overall priority value of the element in the immediate higher level used in the comparisons and h is the number of matrices in the hierarchy.

Second compute the hierarchy dimension as follows:

$$\textbf{Hierarchy dimension n} = \frac{\sum_{i=1}^{h}\left(n_i * g_i^*\right)}{\sum_{i=1}^{h} g_i^*} \tag{9}$$

where n_i is the dimension of matrix i.

The hierarchy is accepted as being fairly consistent if the hierarchy d value is less than or equal to the cut - off value for a matrix of dimension equal to the hierarchy dimension.

SUMMARY.

This paper describe a proposed measure of consistency in the Analytic Hierarchy Process based on multiplicative normalisation. Advantages of the proposed measure includes:

i) independence on the method of estimating weights and the scale used,

ii) it is based on multiplicative normalisation which has been demostrated to have potential of avoiding many of AHP's hortcomings such as the rank reversal, and

iii) the relatively few number and simplicity of computations involved in its use.

The proposed measure of consistency in the AHP was used in determining the consistency of judgement matrices supplied by decision makers as well as the consistency of their hierarchies. Further study is needed to investigate how the proposed measure performs in detecting ordinal inconsistencies in the decision maker's judgement matrices. Furhter it is based on multiplicative normalisation which has been demonstrated to have a potential of avoiding many of AHP's

REFERENCES

1. Saaty, T. L, Multicriteria Decision Making: The Analytic Hierarchy Process, RWS Publications, Pittsburgh, 1990, page 10
2. Islei, G. and Lockett, A. G., Judgemental Modelling Based on Geometric Least Squares, European Journal of Operational Research, vol. 36, 1988, pp. 27 - 35.
3. Golden B. L. and Wang, Q., An Alternative Measure of Consistency, in The Analytic Hierarchy Process: Application and Studies. Golden, B. Wasil, E. and Harker, P. (eds.) Springer-Verlag, Berlin., 1989
4. Barzilai, J., Consistency Measure for Pairwise Comparison Matrices, School of Computer Science, Technical University of Nova Scotia, Halifax, 1996
5. Liang, T. C. and Sheng, C. L., Comments on Saatyís Consistency Ratio Measure and Proposal of a New Detecting Procedure,

Information and Management Sciences, vol. 1 no. 2, 1990 pp. 56 - 68.

6. Barzilai, J. and Golany, B., AHP Rank Reversal, Normalisation and Aggregation Rules, INFOR, vol 32 no. 2, 1994, pp. 57 - 64.
7. Barzilai, J.; Cook, W. D. and Golany, B., Consistent Weights for Judgements Matrices of the Relative Importance of Alternatives, Operations Research Letters, vol. 6 no. 3, 1987, pp. 131 - 134.

The manuscript was received on 31st July 1997 and accepted for publication, after corrections, on 29th October 1997