A Statistical Criterion of Consistency in the Analytic Hierarchy Process

José Antonio Alonso[1] and Mª Teresa Lamata[2]

[1] Dpto. Lenguajes y Sistemas Informáticos. Universidad de Cádiz
Escuela Superior de Ingeniería, 11003-Cádiz, Spain
joseantonio.alonso@uca.es
[2] Dpto de Ciencias de la Computación e I.A, E.T.S de Ingeniería Informática
Universidad de Granada, 18071-Granada, Spain
mtl@decsai.ugr.es

Abstract. In this paper, we present a statistical criterion for accepting/rejecting the pairwise reciprocal comparison matrices in the Analytic Hierarchy Process. We have studied statistically the consistency in random matrices of different sizes. We are not agree with the traditional criterion of accepting matrices due to their inflexibility and because it is too restrictive when the size of the matrix increases. Our system is capable to adapt the acceptance requirements to different scopes and consistency's necessities. The advantages of our consistency system is the introduction of statistical relativity in the acceptance criterion and the simplicity of the used index, the eigenvalue (λ_{max}).

1 Introduction

Over the last three decades, a number of methods have been developed which use pairwise comparisons of the alternatives and criteria for solving discrete alternatives multicriteria decision making (MCDM [1]).

The Analytic Hierarchy Process (AHP) proposed by Saaty [2,3] is a very popular approach to MCDM, that involves qualitative data. It has been applied during the last twenty years in many situations of decision-making.

The AHP has been used on a wide range of applications in a lot of different fields. The method uses a reciprocal decision matrix obtained by pairwise comparisons such that the information is given in a linguistic form.

The method of pairwise comparisons was introduced by Fechner in 1860 [4] and worked out by Thurstone in 1927 [5]. Based in pairwise comparison, Saaty proposes the AHP [2,3] as a method for multicriteria decision-making. It provides a way of breaking down the general method into a hierarchy of sub-problems, which are easier to evaluate.

In the pairwise comparison method, criteria and alternatives, are presented in pairs of one or more referees (e.g., experts or decision-makers). It is necessary to evaluate individual alternatives, deriving weights for the criteria, constructing the overall rating of the alternatives and identifying the best one.

Let us denote the alternatives by $\{A_1, A_2, ..., A_n\}$ (n is the number of compared alternatives), their current weights by $\{w_1, w_2, ..., w_n\}$ and the matrix of the ratios of all weights by $W = [\, w_i / w_j \,]$.

V. Torra et al. (Eds.): MDAI 2005, LNAI 3558, pp. 67–76, 2005.
© Springer-Verlag Berlin Heidelberg 2005

The matrix of pairwise comparisons $A = [\,a_{ij}\,]$ represents the intensities of the expert's preference between individual pairs of alternatives (A_i versus A_j, for all i,j=1,2,..,n). They are chosen usually from a given scale (9,8,…,1/8,1/9). Given n $\{A_1, A_2, ..., A_n\}$ alternatives, a decision maker compares a pair of alternatives for all the possible pairs, n(n-1)/2, and a comparison matrix A is obtained, where the element a_{ij} shows the preference weight of A_i obtained by comparing with A_j.

The a_{ij} elements estimate the ratios w_i / w_j where w is the vector of current weights of the alternative, which is our goal. All the ratios are positive and satisfy the reciprocity property: $a_{ij} = 1/a_{ji}$ ∀i,j=1,2,..,n.

The a_{ij} elements estimate the ratios w_i / w_j where w is the vector of current weights of the alternative, which is our goal. All the ratios are positive and satisfy the reciprocity property: $a_{ij} = 1/a_{ji}$ ∀*i,j*=1,2,..,n.

This paper is divided into four sections. The first one introduces the AHP method. The second one shows the traditional way of measuring consistency and accepting or rejecting matrices in the AHP and the drawbacks we found in this approach. The third section develops our proposal. Our proposal is an alternative and statistical criterion of acceptance or rejection of AHP matrices due to its consistency. Finally, we present the conclusions of the paper.

2 Traditional Criterion of Consistency in the AHP. Drawbacks

The usual method for computing the ranking and weight of alternatives in the AHP is the eigenvector.

In the eigenvector method, the weight vector is the eigenvector corresponding to the maximum eigenvalue "λ_{max}" of the matrix A. According to the Perron-Frobenius Theorem, the eigenvalue "λ_{max}" is positive and real. Furthermore, the vector "w" can be chosen with all positive coordinates. It is a normalized solution of the following equation:

$$A\,w = \lambda_{max}\,w \tag{1}$$

where "λ_{max}" is the largest eigenvalue of the matrix.

The traditional eigenvector method for estimating weights in the Analytic Hierarchy Process yields a way of measuring the consistency of the referee's preferences arranged in the comparison matrix. The consistency index (*CI*) is given by

$$CI = \frac{\lambda_{max} - n}{n - 1} \tag{2}$$

Saaty [2] has shown that if the referee is completely consistent then "$a_{ij} \cdot a_{jk} = a_{ik}$ (∀ i,j,k)", "$\lambda_{max} = n$" and "$CI = 0$". Otherwise,, if the referee is not absolutely consistent "$\lambda_{max} > n$" and Saaty proposes the following index for measuring consistency :

$$CR = CI / RI \tag{3}$$

where '*RI*' is the average value of '*CI*' for a random matrices using the Saaty's scale obtained by Forman[6]. A historical study of several used RIs and a way of estimation this index could be seen in Alonso and Lamata [10]. The main idea is that CR is a normalized value, because is divided by a arithmetic mean of a random matrices consistency indexes (*RI*).

Table 1. Random index for a several matrix dimensions. Forman (17672 to 77847 matrices)

n	1-2	3	4	5	6	7	8	9
RI	0,00	0.52	0.89	1.11	1.25	1.35	1.45	1.49

In the ideal case of total consistency "$\lambda_{max} = n$", the relations between the weights w_i and the judgments a_{ij} will be given by $\frac{w_i}{w_j} = a_{ij}$ for i,j = 1,2,…n.

In this exceptional case the two different matrices, – judgments and weights – are equal. However, it would be unrealistic to require these relations to hold in the general case.

Saaty suggests that a consistency index less or equal to 0.10 indicates that the decision maker has adequately structured the problem in question, but if the consistency index is greater than 0.10 then the response by subject can be considered as random.

But, is this really true? We don't agree with it, and the best way to illustrate it is by giving a pair of examples.

Suppose that two referees R_1 and R_2 express their preferences $P(A_i, A_j)$ about three alternatives $\{A_1, A_2; A_3\}$

$$R_1 \begin{cases} P(A_1, A_2) = 7 \\ P(A_2, A_3) = 1/5 \\ P(A_1, A_3) = 5 \end{cases} \quad \text{and} \quad R_2 \begin{cases} P(A_1, A_2) = 7 \\ P(A_2, A_3) = 1/5 \\ P(A_1, A_3) = 6 \end{cases}$$

We may see that a little difference in the appreciation concerning the preference between A_1 and A_3 yields that the CR (R_1) = 0.0914 (accepted), and CR (R_2) = 0.12 (rejected) to the second subject. We think otherwise that the two set of preferences are nearly the same,

$$P(A_1,A_2) > 1,\ P(A_1,A_3) > 1,\ P(A_3,A_2) > 1$$
$$P(A_1,A_2) \geq \text{Max}\{\ P(A_1,A_3),\ P(A_3,A_2)\ \}$$
$$7 \geq \text{Max}\ \{5, 5\} \text{ for the first case}$$
$$7 \geq \text{Max}\ \{6, 5\} \text{ for the second one}$$

so, we accept the matrix in the first case and the second one will be rejected. We are absolutely disagree with this decision.

As a second example, this matrix

$$\begin{pmatrix} 1 & 5 & 9 \\ 1/5 & 1 & 6 \\ 1/9 & 1/6 & 1 \end{pmatrix}$$

is rejected because, $\lambda_{max} = 3.1632$ and $CR = 0.14 > 0.1$.

In this particular case, it is possible to proof that the transitivity property is not violated.

Therefore: A_1 is preferred (five times) to A_2, A_2 is preferred (six times) to A_3 and A_1 is preferred (nine times) to A_3. Being

$$9 \geq \max\{5, 6\}$$

it is not possible to use a value greater than 9 using the Saaty's scale. In our opinion this matrix must be accepted as a consistent matrix depending on the level of consistency needed.

Seeing these two clear examples, we absolutely disagree with this approach. The problem of accepting/rejecting matrices has been greatly discussed, especially the relation between the consistency and the scale used to represent the decision-maker judgments.

Lane and Verdini (1989) [7] have shown that using a 9-point scale, Saaty's CR threshold is so much restrictive due to the standard deviation of CI for randomly generated matrices is relatively small.

On the other hand Murphy in 1993 [8] have shown that the 9-point scale proposed by Saaty gives results which are outside the accepted consistency when n increases.

Salo and Hämäläinen (1993) [9] have shown that the CR threshold depends on the granularity of the scale, which is being used. Taking into account these ideas we want to introduce a relative and statistical criterion of matrix acceptance. The system can be adapted to whatever scale needed.

Kwiesielewicz [11] introduce the concept of contradictory judgments and Lamata and Pelaez [12,13] have studied different methods to study and improve consistency.

3 A New Statistical Criterion of Matrix Acceptance

The Analytic Hierarchy Process provides the decision-maker an index for measuring the consistency of pairwise reciprocal comparison matrices (CI). Although it is one of the most commonly used methods, it presents some disadvantages. One of these disadvantages is that this index is an absolute index. As we saw in the previous section, this fact causes reasonable consequences little, like accepting or rejecting matrices by minimum differences in the preferences or rejecting matrices that the common sense says to us that they are reasonably consistent.

In this paper we present a new criterion for acceptance and a new index for representing the consistency in pairwise reciprocal comparison matrices. This index and criterion allows the decision-maker to study the consistency of each matrix in a relative way. Using the index that we present, the user can decide about the matrix consistency using not only the matrix entries but also the level of consistency that the decision-maker needs in this particular case.

In this section, we study the AHP consistency problem with statistical criteria. We are going to use as a consistency index the maximum right eigenvalue (λ_{max}) of each studied matrix.

The main idea is that a matrix is consistent (or not) depending on the scope. In different situations, the decision-making could need different levels of consistency and he/she can represent them using percentiles. Therefore, one specific matrix is consistent or it is not (is accepted or not as a consistent matrix) depending on two different factors:

a) An index of consistency (λ_{max}).
b) The level of consistency needed (percentile)

In that case, we can define the consistency of a specific matrix as a Boolean function with two parameters, CR and percentile.

$$F(\lambda_{max}, \text{percentile})$$

- Choosing the percentile by the decision-maker and knowing the number n of alternatives we get a value in Table 3 (λ_{max} T).
- Comparing the matrix CR to the entry λ_{max} T (percentile, n) of the table 3 and if

  $$\lambda_{max} \leq \lambda_{max} T$$

  the matrix is accepted, otherwise

  $$\lambda_{max} > \lambda_{max} T$$

  the matrix is rejected.

We are going to explain the algorithm that we use to generate the λ_{max} values table (λ_{max}T) and that we use to decided if one matrix must be accepted as a consistent one or not.

3.1 Algorithm for Obtaining the Table of λ_{max} Values Depending on Dimension and Percentiles

This algorithm is composed of the following sequence of steps:

1. Generation of the matrices that we are going to study.
2. λ_{max} calculus of each matrix (for all the matrices).
3. Generation of the λ_{max} table (λ_{max} T) depending on dimension and chosen percentiles.

3.1.1 Generation of the Matrices

We have generated 1,200,000 positive reciprocal pairwise comparison matrices (100,000 matrices for each dimension, from 3x3 to 14x14), whose entries were randomly generated using the scale 1/9, 1/8, ...1/2,1,2...8,9 and using a uniform distribution.

```
Algorithm Generation_of_AHP_matrices
   numbermatrices = 100000;
   dimensionfrom = 3;
   dimensionuntil = 14;
for dim= dimensionfrom to  dimensionuntil
   for i=1 to numbermatrices
      mats"dimxdim"(:,:,i)=
                   generates_AHP_single_matrix(dim);
   end
end
End_of_Algorithm Generation_ of_ AHP_ matrices
function  mat = generates_ahp_matrix(dim)
   for i=1 to dim
      for j=1 to dim
         if i>j
            new_value = calculate_AHP_value(random);
```

```
            mat(i,j) = new_value;
            mat(j,i) = 1/new_value;
         end;
      end;
   end;
End_of_function_generates_ahp_matrix
function val = calculate_AHP_value (value)
  Saaty_values=[1/9,1/8,1/7,1/6,1/5,1/4,1/3,1/2,1,2,3
               ,4,5,6,7,8,9]
  ind = Int (value * 17) + 1
  val = Saaty_values[ind]
End_of_function_calculate_AHP_value
```

3.1.2 λ_{max} Calculus of Each Matrix (For All the Matrices)

We are going to calculate the maximum eigenvalue (λ_{max}) of all the matrices (1,200,000) we generated before.

```
Algorithm Calculus_λmax _all_matrices
    numbermatrices = 100000;
    dimensionfrom = 3;
    dimensionuntil = 14;
    for dim= dimensionfrom to  dimensionuntil
       for i=1 to numbermatrices
          LMAX"dimxdim"(i) =   max(eig(mat))
       end
   end
End_of_Algorithm Calculus_λmax _all_Matrices
```

This function uses two Matlab functions, max and eig.

The λ_{max} (3x3) and λ_{max} (4x4) distributions seems to be a Weibull distribution. The rest of distributions are normal distributions. The mean and the standard deviation can be seen in Table 2.

Table 2. On show the mean and deviation of the λ_{max} values

n	3	4	5	6	7	8	9	10	11	12	13	14
$\overline{\lambda}_{max}$	4.051	6.644	9.434	12.241	15.05	17.837	20.598	23.370	26.142	28.903	31.658	34.408
Dstd	1.394	1.883	2.036	2.037	1.985	1.949	1.910	1.871	1.851	1.840	1.811	1.822

As you can see in Fig. 13 the least-squares adjustment line is $\overline{\lambda}_{max}$ = 2.7706 n – 4.356. The x-axis represents the matrix size and the y-axis presents the values of the correspondent values of $\overline{\lambda}_{max}$ (sizes 3 to 14). The correlation coefficient is 0.9999.

3.1.3 Generation of the λ_{max} Table Depending on Dimension and Percentiles

To develop the critical value of λ_{max} that allows us to accept the same percentile of matrices for all values of n, we use the percentile or percentage of the population of λ_{max} and we have used exactly the same scale that Saaty and Forman used to calculate the random index RI, as we present in the algorithm above and the λ_{max} values calculated previously.

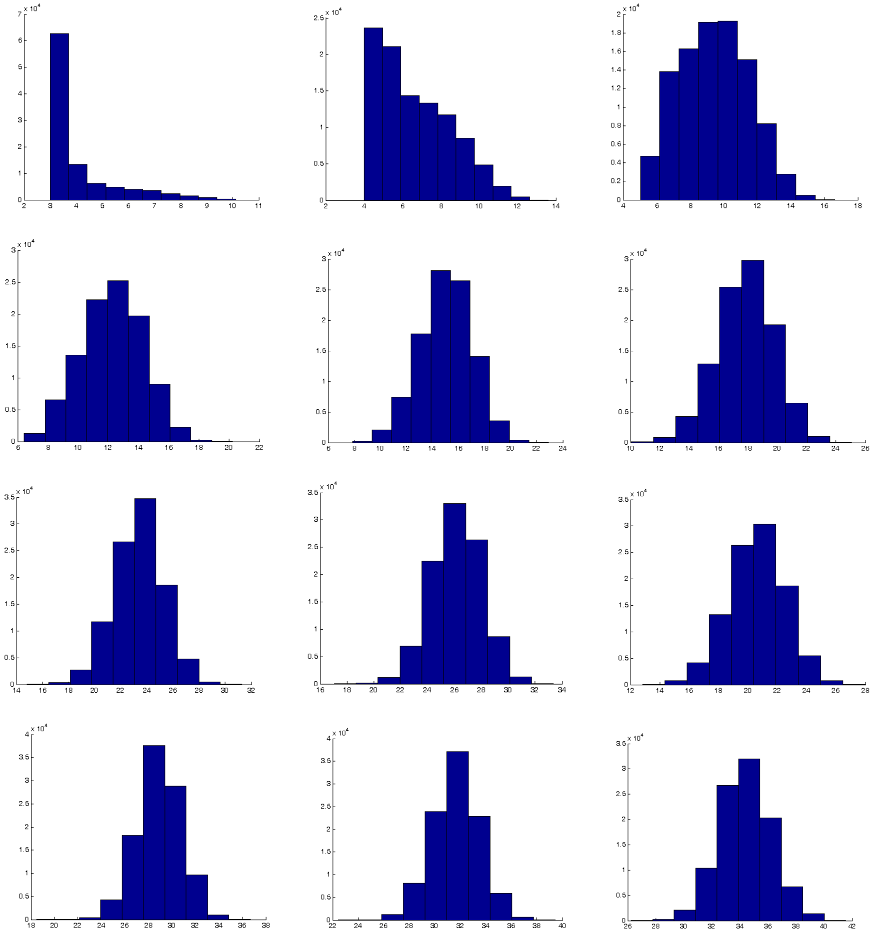

Fig. 1. to Fig 12. Represents the λ_{max} distribution histograms (matrices of dimension from 3x3 to 14x14)

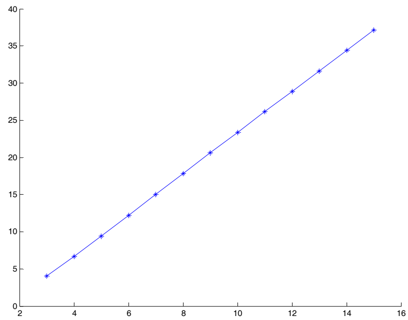

Fig. 13. The different points you can see are corresponding to $\overline{\lambda}_{max}$ of each size (from 3 to 14), and the line is the least-squares adjustment straight line

```
Algorithm Generation_table_LMAX_dimension_percentiles
    number_of_percentiles = 13
    percentiles  =  [0,1,5,10,15,20,25,30,35,40,45,50,100];
    numbermatrices = 100000;
    dimensionfrom = 3;
    dimensionuntil = 14;
    for dim= dimensionfrom to  dimensionuntil
      for p =1 to number_of_percentiles
         LMAXTablep_valuexdim(p,dim)=prctile
             (LMAX"dimxdim"(dim),percentiles(p));
      end
    end
End_of_Algorithm
Generation_table_LMAX_dimension_percentiles
```

For each percentile, p ∈ [0,1], we show the population's value where the (100p)% of the population is on the left and (100 (1-p)%) is on the right. If p is equal to 0, we get the minimum of the population, if p is equal to 1; we obviously get the maximum of the population.

Table 3. On show the λ_{max} values obtained for several percentiles and number of alternatives

	Number of alternatives											
p	3	4	5	6	7	8	9	10	11	12	13	14
Min	3.000	4.000	5.004	6.435	7.876	10.073	12.811	14.818	17.039	18.474	22.415	26.249
0.01	3.000	4.118	5.689	7.703	10.237	13.106	15.975	18.796	21.685	24.488	27.325	30.128
0.05	3.007	4.333	6.217	8.747	11.640	14.535	17.364	20.212	23.095	25.821	28.638	31.441
0.10	3.029	4.518	6.664	9.480	12.425	15.297	18.110	20.956	23.779	26.537	29.326	32.125
0.15	3.053	4.678	7.070	10.023	12.962	15.816	18.614	21.444	24.228	27.006	29.780	32.582
0.20	3.094	4.834	7.455	10.462	13.381	16.217	19.011	21.826	24.580	27.370	30.144	32.936
0.25	3.135	5.004	7.830	10.832	13.731	16.556	19.335	22.145	24.887	27.688	30.444	33.231
0.30	3.197	5.186	8.185	11.160	14.046	16.862	19.631	22.427	25.166	27.976	30.720	33.501
0.35	3.252	5.399	8.516	11.464	14.334	17.140	19.909	22.691	25.439	28.228	30.979	33.753
0.40	3.313	5.650	8.829	11.747	14.597	17.399	20.165	22.942	25.692	28.473	31.219	33.987
0.45	3.367	5.936	9.135	12.022	14.858	17.650	20.413	23.179	25.935	28.705	31.448	34.219
0.50	3.435	6.243	9.433	12.293	15.111	17.897	20.655	23.415	26.170	28.937	31.674	34.446
Max	10.111	13.595	16.615	20.215	22.904	25.074	27.981	31.276	33.373	36.672	39.433	41.551

3.2 Relation Between Our Acceptance System and Saaty's Acceptance System

Using the definition of Consistency Index (2) and the definition of the Consistency Ratio (3) and the traditional acceptance criterion

$$CR = CI / RI < 0.1 \tag{4}$$

and taking into account the definition of RI, as a mean of CI

$$RI = \frac{\overline{\lambda}_{\max} - n}{n-1}$$

we can infer that

$$CR = \frac{\lambda_{\max} - n}{\overline{\lambda}_{\max} - n} < 0.1$$

thus Saaty only accepts matrices as a consistent one, if and only if

$$\lambda_{\max} < n + 0.1\left(\overline{\lambda}_{\max} - n\right).$$

Using the least-squares adjustment straight line we calculated before

$$\overline{\lambda}_{\max} = 2.7706 \text{ n } -4.356 \quad (5)$$

we can conclude that the consistency criterion of Saaty using eigenvalue can be expressed as

$$\lambda_{max} < n + 0.1(17706\text{n } -4.356) \quad (6)$$

and thus

$$\max\lambda_{\max}\ err = 0.1(1.7706\text{n } -4.356) \quad (7)$$

must be the maximum error that Saaty´s accepts in the $\lambda_{\max}$.

Table 4. On show the maximum $\lambda_{\max}$ accepted by Saaty depending on the number of alternatives (3 to 14)

n	3	4	5	6	7	8	9	10	11	12	13	14
λ_{max}	3.095	4.272	5.449	6.626	7.803	8.980	10.157	11.335	12.512	13.689	14.866	16.043

Comparing the values of Table 3 and Table 4 we notice that the criterion of Saaty is too much restrictive.

As we have said the criterion for accepting/rejecting matrices using this table (Table 3) is simple and clear. We accept a matrix as a consistent matrix if and only if its maximum eigenvalue (λ_{max}) is not greater than the value we can get from the table (the row of percentile and column of its dimension). The relativity of this criterion appears when you choose several percentiles for each different situations (or scopes), and its simplicity get using directly the λ_{max}.

It is important notice that, if we use the traditional AHP acceptance criterion, we cannot accept none of them (none of 100,000) of the matrices of dimension greater than 6x6. We are using exactly the same quality data (random matrices), and, in that situation, the traditional matrix acceptation depends not only of the data quality (consistency) but also the dimension of the matrices (number of alternatives concerned). Saaty haven't took into account this fact and the traditional acceptance criterion is absolute and, consequently, too restrictive when the number of alternatives increases.

4 Conclusions

As we have shown, the Saaty's AHP acceptance criterion is absolute and too restrictive when the number of alternatives increases. We have studied statistically the consistency in random matrices of different dimensions. We realized that the value of whatever consistency index depends not only of the consistency of the data (entries of the matrix) but also its dimension.

The AHP criteria haven't took into account this situation and, for that reason, is not appropriate to work with matrices of different dimensions. We took into account this situations, comparing the value of the used consistency index (λ_{max}) for different matrices of the same size.

Our system is able to accept different levels of consistency needed (to adapt the criterion to more or less restrictive situation), and uses a statistical criterion to decide if accepts or not a matrix as a consistent one. Our system compares the matrix level of consistency when the level of consistency of the rest of the matrices of the same dimension. This statistical criterion, taking into account the dimension of the matrices, and the different level of consistency needed, offer clear advantages compared with the traditional system, and we use a more simple consistency index than Saaty´s, the maximum eigenvalue of each matrix (λ_{max}).

Acknowledgments

This work was partially supported by the DGICYT under project TIC2002-03411.

References

1. Keeney, R., Raiffa, H.: Decisions with Multiple Objectives; Preferences and Values Tradeoffs. Wiley. New York (1976)
2. Saaty, T.L.: The Analytic Hierarchy Process. New York (1980)
3. Saaty, T.L.: Fundamentals of Decision Making and Priority Theory with the Analytic Hierarchy Process. RWS Publications Pittsburgh (1994)
4. Fechner,G.T.: Elements of Psychophysics, Volume 1, Holt, Rinehart & Winston, New York, (1965); translation by H.E.Adler of Elemente der Psychophysik, Breitkopf und Hârtel, Leipzig (1980)
5. Thurstone, L.L.: A law of comparative judgments. Psychological Reviews 34, (1927) 273-286.
6. Forman, E.H.: Random indices for Incomplete Pairwise Comparison Matrices. European Journal of Operational Research 48, (1990) 153-155.
7. Lane, E.F., Verdini, W.A.: A consistency test for AHP decision makers. Decision Science 20, (1989) 575-590.
8. Murphy, C.K.: Limits of the Analytical Hierarchy Process from its consistency index. European Journal of Operational Research 65, (1993) 138-139
9. Salo A.A., Hämäläinen R.P.: On the measurement of preferences in the AHP", Research Report, Systems Analysis Laboratory, Helsinki University of Technology (1993)
10. Alonso. J.A, Lamata M.T.: Estimation of the Random Index in the Analytic Hierarchy Process. Proceedings of Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU'04) Perugia Vol I, (2004) 317-322.
11. Kwiesielewicz. M., van Uden.E.: Inconsistent and contradictory judgements in pairwise comparison method in the AHP Computers & Operations Research 31, (2004) 713–719
12. Pelaez.J.I., Lamata.M.T.: A New Measure of Consistency for Positive Reciprocal Matrices. Computers and Mathematics with Applications 46, (2003) 1839-1845
13. Lamata.M.T., Pelaez.J.I.: A method for improving the consistency of judgements. International Journal of Uncertainty Fuzzyness and Knowledge-based Systems 10 (2002) 677-686