A Chi-square-based inconsistency index for pairwise comparison matrices

Michele Fedrizzi[†] , Fabio Ferrari

[†] corresponding author, Department of Industrial Engineering

University of Trento, Via Sommarive 9, I-38123 Trento, Italy

e–mail: `michele.fedrizzi@unitn.it`

Abstract

In this paper we introduce a new method for evaluating the inconsistency level of a pairwise comparison matrix. The classical Chi-square index suggests an interesting formal similarity for a consistent pairwise comparison matrix and, as a consequence, a method for measuring the relative deviation of the elicited preferences from a set of consistent preferences defined on the basis of the similarity mentioned above. Contrary to some previously introduced Chi–square–based approaches, no optimization problems are involved. We verify that the new index satisfies some recently introduced characterizing properties of inconsistency indices. Then, by means of numerical simulations, we compare our index with some other well-known inconsistency indices and we focus, in particular, on the comparison with Saaty's consistency index. We discuss some numerical results showing that the new index is closely related with Saaty's one but it is more stable with respect to the number of alternatives.

Keywords: analytic hierarchy process, pairwise comparison matrices, consistency indices.

1 Introduction

Pairwise comparison is a well established and popular method for ranking alternatives. The well-known Analytic Hierarchy Process (AHP) by T. Saaty

(Saaty, 1977), in particular, is based on pairwise comparisons on a ratio scale and a hierarchical structure of the decisional problem. The eigenvector method is used to derive the weight vector for the alternatives from each Pairwise Comparison Matrix $\mathbf{A} = (a_{ij})$, PCM in the following.

Measuring the inconsistency of the elicited preferences is a crucial point of the methodology, since consistency of the decision maker's judgements is closely related to their reliability and, as a consequence, to the reliability of the obtained weights. In order to evaluate how far from consistency a PCM is, Saaty proposed, in his seminal paper (Saaty, 1977), the first and still most popular inconsistency index. Saaty's index is based on the principal eigenvalue of the PCM and refers to the Frobenius theorem. Nevertheless, Saaty's approach was criticized by many authors (Peláez and Lamata, 2003; Barzilai, 2005; Bana e Costa and Vansnick, 2008; Bozóki and Rapcsák, 2008; Grzybowski, 2012, 2016) and numerous different approaches to inconsistency evaluation have been proposed in the literature with many different definitions for inconsistency indices. The majority of the indices introduced so far are closely related with a particular prioritization procedure and reflect the global deviation of the elicited preferences a_{ij} from the estimated ratio of weights w_i/w_j. A comprehensive discussion of all the already known indices is beyond the scope of this paper. The reader may refer to the introduction of (Grzybowski, 2016) for a concise and up-to-date description. The relevant attention and interest of the scientific community on this topic evidences that no satisfactory agreement has yet been reached on inconsistency evaluation, but a deeper understanding on the problem is in progress.

We mention some few recent contributions that introduce new points of view in the debate. Temesi (Temesi, 2011) introduced the distinction between consistent and error-free matrix. Grzybowski (Grzybowski, 2016) studied the relationship between inconsistency indices and the quality of priority vector estimation. Cavallo et al. introduced a consistency index in the general algebraic framework of an abelian group, see (Cavallo et al, 2012) and other preceding papers by the same authors. Bozóki and Rapcsák (Bozóki and Rapcsák, 2008), Brunelli et al. (Brunelli et al, 2013a), Kazibudzki (Kazibudzki, 2016), and Grzybowski (Grzybowski, 2016) proposed numerical studies aimed at investigating and comparing different inconsistency indices. Sets of axioms for inconsistency indices have been proposed and discussed by Koczkodaj and Szwarc (Koczkodaj and Szwarc, 2014), Brunelli and Fedrizzi (Brunelli and Fedrizzi, 2015), Brunelli (Brunelli, 2017) and Mazurek (Mazurek, 2017). A set of axioms characterizing an incon-

sistency ranking was proposed by Csató (Csató, 2016). It is interesting to observe that the studies mentioned above show that some indices behave very differently from each other, whereas some other indices, even if defined starting from completely different points of view, share the same properties and a very similar numerical behavior. This latter case, in our view, strengthens the reliability of the involved indices and is not at all trivial. A coherent behavior of seemingly different indices means that they are mutually supportive as they can bring separate but confirming evidence on the consistency evaluation.

Our proposal is motivated by the preceding remarks, as we start from a point of view that is different from the previous approaches. We are confident that our work will help a better understanding of some aspects of inconsistency evaluation. The starting point of our paper is a technical remark pointed out by Lipovetsky and Conklin in (Lipovetsky and Conklin, 2002). These authors proposed a new interpretation of a pairwise comparison matrix with the aim of identifying what they called 'unusual and false observations'. In our opinion, the remark of Lipovetsky and Conklin can also lead to an interesting new approach to inconsistency evaluation of pairwise comparison matrices. Unlike Lipovetsky and Conklin's approach, our proposal describes some results we obtained by developing their formal interpretation without assuming any statistical meaning or acknowledgment. We define our inconsistency index $I_{\chi^2}(\mathbf{A})$ of a pairwise comparison matrix $\mathbf{A}$ by means of the corresponding χ^2 value associated to $\mathbf{A}$ by means of the formal interpretation introduced in (Lipovetsky and Conklin, 2002). Then, we study index $I_{\chi^2}(\mathbf{A})$ both from the theoretical point of view and by means of numerical simulations. The paper is organized as follows. Section 2 describes the necessary basic notions and notations on pairwise comparison matrices, AHP and inconsistency indices. In Section 3 the new inconsistency index is defined and commented. Then, other known indices referring to Chi–square are briefly recalled in order to highlight the differences of the various approaches. In Section 4, some relevant properties of our index are studied referring to six axioms recently introduced in (Brunelli and Fedrizzi, 2015; Brunelli, 2017). Numerical results are given in Section 5, where we study our index in comparison with Saaty's consistency index as well as with some other inconsistecy indices. Some final remarks are discussed in Section 6. We note that an index which estimates the deviation from consistency is often referred to as 'consistency index'. Nevertheless, in this paper the term 'inconsistency index' is used, since we consider it more suitable to measure the

deviation from consistency, taking into account that consistency is always identified by a single optimal value of the index.

2   Background

Given a set of alternatives $X = \{x_1, \ldots, x_n\}$ $(n \geq 2)$, a pairwise comparison matrix $\mathbf{A} = (a_{ij})_{n \times n}$ (PCM in the following) is a positive and reciprocal matrix of order n, i.e. $a_{ij} > 0$, $a_{ij}a_{ji} = 1$, $\forall i, j$. The positive real value a_{ij} is a numerical estimation of the ratio w_i/w_j, where w_i and w_j are the weights of x_i and x_j respectively (Saaty, 1977). A pairwise comparison matrix is defined to be *consistent* if and only if the following transitivity condition holds:

$$a_{ik} = a_{ij}a_{jk} \quad \forall i, j, k. \tag{1}$$

If and only if $\mathbf{A}$ is consistent, then there exists a priority (or weight) vector $\mathbf{w} = (w_1, \ldots, w_n)$ such that

$$a_{ij} = \frac{w_i}{w_j} \quad \forall i, j. \tag{2}$$

If $\mathbf{A}$ is consistent, then the components of the priority vector $\mathbf{w}$ can be obtained by using the eigenvector method (Saaty, 1977) or, more simply, by using the geometric mean method (Crawford and Williams, 1985):

$$w_i = \left(\prod_{j=1}^{n} a_{ij} \right)^{\frac{1}{n}} \quad \forall i. \tag{3}$$

Note that some authors pointed out the difference between a 'consistent PCM' and an 'error–free PCM' (Temesi, 2011; Grzybowski, 2016). Normalization condition $\sum_{i=1}^{n} w_i = 1$ is often imposed in order to have uniqueness of the weight vector. Since it is known that 'in making paired comparisons people do not have the intrinsic logical ability to always be consistent' (Saaty, 1994a), a violation of (1) is to some extent necessarily accepted. Nevertheless, coherent judgements are clearly considered more preferable and reliable than contradictory ones and, thus, a satisfactory consistency level has commonly been regarded as a desirable property. Therefore, a correct inconsistency evaluation is regarded as a crucial task, and several indices have been proposed in order to quantify the deviation from the condition of full consistency

(1) or (2). We omit, for brevity, the description of the numerous inconsistency indices proposed in literature (Saaty, 1977; Crawford and Williams, 1985; Golden and Wang, 1989; Koczkodaj, 1993; Duszak et al, 1994; Salo and Hämäläinen, 1995; Barzilai, 1998; Shiraishi et al, 1998; Fedrizzi et al, 2002; Peláez and Lamata, 2003; Stein and Mizzi, 2007; Fedrizzi and Brunelli, 2009; Cavallo and D'apuzzo, 2009; Ramík and Korviny, 2010; Cavallo et al, 2012; Wu and Xu, 2012; Brunelli et al, 2013b; Kou and Lin, 2014; Grzybowski, 2016). The reader can see the references. We only briefly recall the first introduced and best known index, i.e. the 'Consistency Index' by T. Saaty (Saaty, 1977),

$$CI(\mathbf{A}) = \frac{\lambda_{max} - n}{n - 1}. \tag{4}$$

In (4), $\lambda_{max} \geq n$ denotes the principal right eigenvalue of $\mathbf{A}$. It was proved by Saaty that $\lambda_{max} = n$ if and only if $\mathbf{A}$ is consistent, whereas all the remaining $n - 1$ eigenvalues equal to zero. As a consequence, the rank of $\mathbf{A}$ is one and all columns (rows) are proportional. Index (4) has been normalized by Saaty as follows, thus obtaining the 'Consistency Ratio' CR ,

$$CR(\mathbf{A}) = \frac{CI(\mathbf{A})}{RI} . \tag{5}$$

RI denotes the 'random index', which is the mean CI computed over a large number of randomly generated pairwise comparison matrices of the same order of $\mathbf{A}$. Saaty computed the random index in his seminal paper (Saaty, 1977) using 50 random matrices for each n. Then, Alonso and Lamata refined his results (Alonso and Lamata, 2003). Saaty proposes to accept a pairwise comparison matrix as sufficiently consistent if the consistency ratio CR is below the threshold 0.1.

Among the most popular indices, we also cite the Geometric Consistency index (Crawford and Williams, 1985) and the Koczkodaj index (Koczkodaj, 1993).

3   Chi–square based inconsistency index

It was proved in (Lipovetsky and Conklin, 2002) that the entries of a consistent matrix $\mathbf{W} = (w_{ij})$ can be written in the form

$$w_{ij} = \frac{(\sum_{k=1}^{n} w_{ik})(\sum_{h=1}^{n} w_{hj})}{\sum_{h=1}^{n} \sum_{k=1}^{n} w_{hk}}. \tag{6}$$

Therefore, Lipovetsky and Conklin (Lipovetsky and Conklin, 2002) pointed out that matrix $\mathbf{W}$ can formally be interpreted as a contingency table. Nevertheless, we do not assume any statistical setting and we simply consider (6) as one of the possible characterizations of a consistent PCM.

Let us now assume that a PCM $\mathbf{A} = (a_{ij})$ is obtained by means of the subjective judgements of an expert. Matrix $\mathbf{A}$ is, in general, inconsistent and we will refer to this matrix as to an 'empirical' matrix. Table 1 represents such a matrix by evidencing the margins which refer to expression (6). In fact, although property (6) is not satisfied by an inconsistent matrix $\mathbf{A}$, we can associate to $\mathbf{A}$ a consistent matrix $\mathbf{E} = (e_{ij})$ which elements e_{ij} are defined similarly to (6),

$$e_{ij} = \frac{(\sum_{k=1}^{n} a_{ik})(\sum_{h=1}^{n} a_{hj})}{\sum_{h=1}^{n} \sum_{k=1}^{n} a_{hk}}. \tag{7}$$

If the values e_{ij} are close to the empirical data a_{ij}, then matrix $\mathbf{A} = (a_{ij})$ is close to consistency. As noted in the previous section, a pairwise comparison matrix, i.e. a positive reciprocal matrix, is consistent if and only if all its columns (rows) are proportional. Matrix $\mathbf{E} = (e_{ij})$ clearly satisfies this requirement.

	1	2	.	.	.	.	n	row total
1	a_{11}	a_{12}	.	.	.	.	a_{1n}	$\sum_{j=1}^{n} a_{1j}$
2	a_{21}	a_{22}	.	.	.	.	a_{2n}	$\sum_{j=1}^{n} a_{2j}$
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
n	a_{n1}	a_{n2}	.	.	.	.	a_{nn}	$\sum_{j=1}^{n} a_{nj}$
column total	$\sum_{i=1}^{n} a_{i1}$	$\sum_{i=1}^{n} a_{i2}$	.	.	.	.	$\sum_{i=1}^{n} a_{in}$	$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}$

Table 1: Empirical pairwise comparison matrix

The previous remarks suggest that the χ^2 index is a suitable tool to evaluate the deviation of empirical data a_{ij} from the associated consistent ones e_{ij}. Therefore, we define the χ^2 index of $\mathbf{A}$ as

$$\chi^2(\mathbf{A}) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{(a_{ij} - e_{ij})^2}{e_{ij}} \ . \tag{8}$$

6

We remark again that expression (8) is only formally taken from the known χ^2 formula, but we do not assume here any statistical setting. Similarly to (4), index (8) depends on the order n of matrix $\mathbf{A}$. Then, following Saaty, we define the $\chi^2 - RandomIndex$ $RI_n(\chi^2)$ as the mean value of $\chi^2(\mathbf{A})$ computed over a large number of randomly generated pairwise comparison matrices of the same order n of $\mathbf{A}$. Values of $RI_n(\chi^2)$ are reported in Table 2 for n from 3 to 8.

n	$RI_n(\chi^2)$
3	6.552
4	21.022
5	43.150
6	72.258
7	108.759
8	151.197

Table 2: Random Index $RI_n(\chi^2)$

The values in Table 2 are obtained by means of 50,000 randomly generated PCMs using Saaty's scale, i.e. $a_{ij} \in \{1/9, ..., 9\}$. We finally define the normalized χ^2–based inconsistency index of a matrix $\mathbf{A}$ of order n as

$$I_{\chi^2}(\mathbf{A}) = \frac{\chi^2(\mathbf{A})}{RI_n(\chi^2)} \ . \tag{9}$$

and we call it 'χ^2–based inconsistency index'.

Example 1. *Let us consider a pairwise comparison matrix $\mathbf{A} = (a_{ij})$ of order 4,*

$$\mathbf{A} = \begin{pmatrix} 1 & 4 & 5 & 7 \\ 0.25 & 1 & 2 & 4 \\ 0.2 & 0.5 & 1 & 2 \\ 0.143 & 0.25 & 0.5 & 1 \end{pmatrix} . \tag{10}$$

The Chi-square value (8) for (10) is $\chi^2(\mathbf{A}) = 0.7399$ and the normalized index (9) is $I_{\chi^2}(\mathbf{A}) = 0.0352$. Numerical simulations presented in Section 5, will show that such value can classify matrix (10) as satisfactorily consistent. In order to make a comparison, we compute for (10) also the Saaty's Consistency Ratio (5). Since it is $CR(\mathbf{A}) = 0.0246$, matrix $\mathbf{A}$ is considered satisfactorily consistent by both indices.

3.1 Other Chi–square related inconsistency indices

Other methods referring to Chi squares have been proposed in order to derive weights from a pairwise comparison matrix and to evaluate its inconsistency. They are meaningful and correctly justifiable, but they are intrinsically different from our approach. We briefly describe some of these methods, while the reader can find details in the references. The first four methods described in the following are based on the optimization of a deviation function. Jensen (Jensen, 1983) proposed to determine the weight vector $\mathbf{w} = (w_1, ..., w_n)$ by solving the optimization problem

$$\min \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\left(a_{ij} - \frac{w_i}{w_j}\right)^2}{\frac{w_i}{w_j}} \tag{11}$$

$$\sum_{j=1}^{n} w_j = 1 , \quad w_j > 0, \quad j = 1, ..., n ,$$

where the value of the objective function in (11) computed in the optimal solution of the problem is taken as an inconsistency index of matrix (a_{ij}). Jensen's method was extended by Wang et al. in (Wang et al, 2007) by allowing the preferences to be expressed both in the multiplicative (a_{ij}) and in the additive (r_{ij}) representation and by considering a group of $k = 1, ..., m$ decision makers with possibly different a priori importance weights h_k. Then, the weight vector $\mathbf{w} = (w_1, ..., w_n)$ is obtained by solving the optimization problem

$$\min \left(\sum_{k=1}^{m_1} \sum_{i=1}^{n} \sum_{j=1}^{n} h_k \frac{\left(a_{ij}^{(k)} - \frac{w_i}{w_j}\right)^2}{\frac{w_i}{w_j}} + \right. \tag{12}$$

$$\left. \sum_{k=m_1+1}^{m} \sum_{i=1}^{n} \sum_{j=1}^{n} h_k \frac{\left(r_{ij}^{(k)} - \frac{w_i}{w_i+w_j}\right)^2}{\frac{w_i}{w_i+w_j}} \right)$$

$$\sum_{j=1}^{n} w_j = 1 , \quad w_j > 0, \quad j = 1, ..., n ,$$

see reference for details. Z.S. Xu (Xu, 2000) proposed a generalization of Jensen's method by considering the following optimization problem

$$\min \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{a_{ij}^{\alpha} - \left(\frac{w_i}{w_j}\right)^{\alpha}}{\left(\frac{w_i}{w_j}\right)^{\alpha}} \tag{13}$$

$$\sum_{j=1}^{n} w_j = 1 , \quad w_j > 0, \quad j = 1, ..., n .$$

Recently, Y. Xu et al. proposed a method for priority derivation in the framework of group decision making (Xu et al, 2015). This method is similar to Jensen's one, but refers to incomplete reciprocal preference relations. The approach proposed in (Lin et al, 2013) by Lin et al. is different from the four described above since it is based on statistical inference. The authors estimate the sample variance of the errors between the elicited preferences a_{ij} and the expected ones $\frac{w_i}{w_j}$ with a maximum likelihood test.

4 Properties of index $I_{\chi^2}(\mathbf{A})$

In this section we consider a set of six properties introduced in (Brunelli and Fedrizzi, 2015) and in (Brunelli, 2017) in order to characterize an inconsistency index and we study the satisfaction of these properties by the index I_{χ^2}. For sake of brevity, we recall only briefly the six properties. The reader can refer to (Brunelli and Fedrizzi, 2015; Brunelli, 2017) for a more detailed description.

The first characterizing property (P1) requires that there exists a unique real number, say ν, such that the inconsistency index of a matrix is equal to ν if and only if the matrix is consistent. As remarked in Section 3, it is $a_{ij} = e_{ij} \; \forall i, j$ if and only if $\mathbf{A}$ is consistent. Then, index $I_{\chi^2}(\mathbf{A})$ given by (9) is null if and only if $\mathbf{A}$ is consistent. Therefore, $I_{\chi^2}(\mathbf{A})$ satisfies (P1) with $\nu = 0$.

The second characterizing property (P2) requires that an inconsistency index should be independent from the order of the alternatives. More formally, the inconsistency of $\mathbf{A}$ must be equal to the inconsistency of $\mathbf{P}\mathbf{A}\mathbf{P}^T$ for any permutation matrix $\mathbf{P}$. From (8) and (9) it follows directly that $I_{\chi^2}(\mathbf{A})$ satisfies also (P2).

The idea underlying property (P3) is the following: if inconsistent preferences are intensified, then a better value of an inconsistency index cannot be obtained. By 'preference intensification' we mean going farther from complete indifference $a_{ij} = 1 \; \forall i, j$, which leads clearly to a fully consistent matrix. Going farther from this uniformity means having stronger judgments and this should *not* make their possible inconsistency *less* evident. The only transformation $a_{ij} \to f(a_{ij})$ which can intensify preferences and preserve the necessary reciprocity structure, namely $a_{ij}a_{ji} = 1, \; \forall i, j$, is

$$f(a_{ij}) = (a_{ij})^k, \;\; k > 1. \tag{14}$$

This properties was tested by means of numerical simulations and it emerged that (P3) was always satisfied by $I_{\chi^2}(\mathbf{A})$. The simulation framework is the same described in Sections 5.1 and 5.2. More precisely, we tested property (P3) for $I_{\chi^2}(\mathbf{A})$ on two types of PCMs. The first type is randomly generated PCMs. The second type is obtained by starting with consistent PCMs and by applying a perturbation which intensity depends on a parameter σ. We considered $n = 3, ..., 8$ and numerous different values of σ. See Sections 5.1 and 5.2 for more details.

The property (P4) is based on the idea that an inconsistency index is non–decreasing with respect to an elementary modification of a consistent matrix. More formally, given a consistent pairwise comparison matrix $\mathbf{A}$, let us modify the single entry a_{pq} and, accordingly, its reciprocal a_{qp}. The property (P4) requires that the larger the change of a_{pq} from its consistent value, the more inconsistent will be the obtained matrix. We tested (P4) numerically for $I_{\chi^2}(\mathbf{A})$ on the same set of matrices as for property (P3), evidencing that $I_{\chi^2}(\mathbf{A})$ satisfies (P4) too.

Property (P5) requires the continuity of an inconsistency index. Continuity of $I_{\chi^2}(\mathbf{A})$ directly follows from continuity of the functions involved in its definition, see (8) and (9).

The last property, (P6), was recently introduced in (Brunelli, 2017) in order to complete the set of the first five properties introduced in (Brunelli and Fedrizzi, 2015). Property (P6) requires the invariance under inversion of preferences. More formally, the inconsistency of a PCM $\mathbf{A}$ must be equal to the inconsistency of its transpose $\mathbf{A}^T$. From the expressions of (8)and (9), it is straightforward to prove that $I_{\chi^2}(\mathbf{A}) = I_{\chi^2}(\mathbf{A}^T)$. Therefore, I_{χ^2} satisfies (P6).

To summarize, it is possible to prove that $I_{\chi^2}(\mathbf{A})$ satisfies properties (P1), (P2), (P5) and (P6), whereas properties (P3) and (P4) have been verified by means of numerical simulations but not formally proved. The satisfaction of (P3) and (P4) remains, therefore, a conjecture.

5 Numerical results

In this section we present two numerical studies in order to compare the new index I_{χ^2} with some other well-known inconsistency indices. Other interesting numerical comparisons between inconsistency indices are proposed in (Choo and Wedley, 2004; Lin, 2007; Bozóki and Rapcsák, 2008; Grzybowski,

2012, 2016). First, in Subsection 5.1, we follow an approach similar to that of Golden and Wang (Golden and Wang, 1989) in order to investigate the agreement between I_{χ^2} and Saaty's consistency ratio CR. The problem of determining a rigorously justified acceptance threshold for the various inconsistency indices is an important and non easy question that has been investigated by several researches. For a deeper discussion on the topic, see (Bozóki and Rapcsák, 2008; Grzybowski, 2012; Dijkstra, 2013; Grzybowski, 2016). We are aware of the fact that the acceptance threshold 0.1 fixed by Saaty has no formal justification, since it is purely heuristic. In the following, we will refer to this frequently adopted threshold only for comparison purpose. Then, in Subsection 5.2, we select four other inconsistency indices among the most relevant ones and we investigate numerically their relationships with index I_{χ^2}. To this aim, we use three different statistical tools: scatter plots, Pearson's correlation coefficient (Pearson, 1920) and Spearman's rank correlation coefficient (Spearman, 1904; Snedecor and Cochran, 1980).

5.1 Index $I_{\chi^2}(\mathbf{A})$ versus Saaty's Consistency Ratio

In order to compare our index I_{χ^2} with Saaty's CR (5), we first constructed the numerical data set with a method similar to the one proposed in (Brunelli et al, 2013a).

- We generated a set $S = \{\mathbf{B}_1, ..., \mathbf{B}_N\}$ with $N = 10,000$ PCMs of order n obtained by means of a random perturbation on consistent PCMs. More precisely, the following procedure was repeated $N = 10,000$ times: first, a consistent PCM $\mathbf{B} = (b_{ij})$ is constructed by setting $(b_{ij}) = \left(\frac{w_i}{w_j}\right)$, where $(w_1, ..., w_n)$ is a randomly generated vector with $w_i \in [1, 9]$, so that $b_{ij} \in [1/9, 9]$. Then, each consistent PCM is modified by means of a random perturbation on single elements above the diagonal $b_{ij} \to b_{ij}e^{\beta}$, where β is a random variable with normal distribution, $\beta \sim N(0, \sigma)$. The elements below the diagonal of the PCM are modified accordingly to preserve reciprocity, $b_{ji} = 1/b_{ij}$.

- We computed the inconsistency indices $I_{\chi^2}(\mathbf{B}_j)$ and $CR(\mathbf{B}_j)$ for every PCM $\mathbf{B}_j \in S$.

- We repeated the points described above for $n = 3, ..., 8$ and for different values of σ. The larger is the value of σ, the larger is the mean deviation

from the initial consistent matrix and, as a consequence, the more inconsistent is, on average, the obtained PCM.

For sake of space, we report in Table 3 the results of a single significant example of the performed numerical simulations. Upon request, we may provide the interested reader with other similar results in tabular form. Table 3 reports the acceptance or rejection frequencies for $10,000$ perturbed matrices with $\sigma = 0.4$ as for their satisfaction of the consistency thresholds. For both indices (5) and (9) the frequently adopted Saaty's threshold 0.10 is assumed. Following the well-known approach of Golden and Wang (Golden and Wang, 1989), in the columns of Table 3 the following results are reported.

- Column F(A,A) contains the number of PCMs, out of the original $10,000$, that satisfy both the consistency threshold of CR and I_{χ^2}, so that they are classified as sufficiently consistent by the two indices.

- Column F(A,R) contains the number of PCMs, out of the original $10,000$, that satisfy the consistency threshold of CR but not the one of I_{χ^2}, so that they are classified as sufficiently consistent by CR but rejected by I_{χ^2}.

- Column F(R,A) contains the number of PCMs, out of the original $10,000$, that satisfy the consistency threshold of I_{χ^2} but not the one of CR, so that they are rejected by CR but classified as sufficiently consistent by I_{χ^2}.

- Column F(R,R) contains the number of PCMs, out of the original $10,000$, that do not satisfy the consistency threshold of CR neither the one of I_{χ^2}, so that they are rejected by both indices.

- Column F(A,*) contains the total number of PCMs, out of the original $10,000$, that satisfy the consistency threshold of CR, no matter the result for index I_{χ^2}.

- Column F(*,A) contains the total number of PCMs, out of the original $10,000$, that satisfy the consistency threshold of I_{χ^2}, no matter the result for index CR.

- Column 'agreement' contains the total number of PCMs, out of the original $10,000$, that are classified in the same way by CR and I_{χ^2}, that is F(A,A) + F(R,R).

n	F(A,A)	F(A,R)	F(R,A)	F(R,R)	F(A,*)	F(*,A)	agreement
3	8361	0	669	970	8361	9030	9331
4	9059	30	506	405	9089	9565	9464
5	9547	43	245	165	9590	9792	9712
6	9833	36	76	55	9869	9909	9888
7	9921	28	36	15	9949	9957	9936
8	9955	20	15	10	9975	9970	9965

Table 3: Comparing acceptance/rejection frequencies of CR and I_{χ^2}

Let us make some observations and comments from the results in Table 3:

- There is a good agreement between CR and I_{χ^2} in classifying consistency of PCMs, as it emerges from the data in the last column of the table. Note that other known indices behaves differently from CR (Brunelli et al, 2013a), whereas other indices show even a more closer relationship (Grzybowski, 2016).

- The behavior of I_{χ^2} is more stable than CR with respect to n: the coefficient of variation of elements in column F(A,*) is 0.067 whereas the corresponding value for column F(*,A) is 0.037. From this point of view, index I_{χ^2} is more fair than CR in classifying consistency of PCMs with respect to different dimensions n. To support the previous findings, we report in Table 4 the mean values μ_{CR}, $\mu_{I_{\chi^2}}$ and the standard deviations σ_{CR}, $\sigma_{I_{\chi^2}}$ of CR and I_{χ^2} respectively, as obtained from the data set S described above.

n	μ_{CR}	σ_{CR}	$\mu_{I_{\chi^2}}$	$\sigma_{I_{\chi^2}}$
4	0.0462	0.0382	0.0357	0.0306
5	0.0443	0.0266	0.0361	0.0230
6	0.0444	0.0204	0.0379	0.0200
7	0.0437	0.0166	0.0378	0.0160
8	0.0439	0.0138	0.0393	0.0142

Table 4: Mean and standard deviation of CR and I_{χ^2} in set S

The same conclusion on the stability with respect to n holds also for different values of σ, that is for different inconsistency levels.

- An interesting finding emerges from the comparison of the results in Table 3 with other similar results obtained with different values of σ.

13

It concerns a shortcoming of the CR which may occur for small values of n, in particular for $n = 3$. It is known that some 3×3 matrices with apparently large inconsistency may have a CR below the threshold 0.1 and are therefore accepted as sufficiently consistent. In fact, Saaty himself refined the 0.1 rule by suggesting 0.05 threshold for 3×3 PCMs and 0.08 threshold for 4×4 PCMs (Saaty, 1994b). This remark induced the common opinion that CR is less demanding for PCMs with small n than it is for large n. From our study it emerges that this is not true in general, since the behavior depends on the set of matrices which is considered. More precisely, this behavior is limited to the case of matrices with large inconsistency, i.e. to the case of large values of σ. In this case, data show that the percentage of matrices accepted as sufficiently consistent decreases as n increases. On the contrary, in the case of nearly consistent matrices, i.e. small values of σ, the percentage of matrices accepted as sufficiently consistent increases as n increases. As an example, see column F(A,*) of Table 3.

5.2 Numerical comparison with other inconsistency indices

Since the seminal work of Saaty in the 70's, a large number of inconsistency indices has been proposed in the literature. Recent studies were devoted to compare numerically some of these indices (Bozóki and Rapcsák, 2008; Grzybowski, 2012; Brunelli et al, 2013a; Kazibudzki, 2016). It is beyond the scope of this paper to present a detailed comparison between our new proposed index $I_{\chi^2}(\mathbf{A})$ and all the already known inconsistency indices. Therefore, we chose only few of them, selected among the most relevant ones, in order to have the possibility of making some comments and remarks on numerical comparison examples. In particular, we selected Saaty's consistency ratio CR, the index CI^* of determinants by Peláez and Lamata (Peláez and Lamata, 2003), the index GW by Golden and Wang (Golden and Wang, 1989) and the geometric consistency index GCI by Crawford and Williams (Crawford and Williams, 1985; Aguarón and Moreno-Jiménez, 2003). The detailed definitions of the cited indices can be found in the corresponding references.

As in (Brunelli et al, 2013a), we consider two classes of matrices: consistent PCMs perturbed with small random noise, like in the previous subsec-

tion, and randomly generated PCMs. Clearly, the former class emulates a real world case, whereas the latter is a neutral benchmark case.

In both cases, we first compare I_{χ^2} with the other indices graphically, by means of scatter plots. Then, we compute Pearson's correlation coefficient between I_{χ^2} and each one of the selected indices in order to highlight the linear correlation between them. Finally, by means of Spearman's rank correlation coefficient, we study the comonotonicity of the association between I_{χ^2} and each one of the selected indices.

Case 1: consistent PCMs perturbed with small random noise We

consider a set of PCMs generated as described in Subsection 5.1, with $n = 6$ and $\sigma = 0.4$. Nevertheless, in order to have a better graphical rendering, the number of PCMs is reduced to $1,000$. We chose to generate PCMs perturbed by $\sigma = 0.4$ since this corresponds to a reasonably low level of inconsistency, as could be found in a real-world case (see Table 3). The graphical comparisons between I_{χ^2} and the other indices are presented in Figure 1. Let us describe, as an example, the scatter plot in Figure 1a. Each point in the plot represents a PCM, say $\mathbf{B}_j$, the coordinates being respectively the inconsistency values $I_{\chi^2}(\mathbf{B}_j)$ and $CR(\mathbf{B}_j)$ for that matrix. The plots in Figures 1b, 1c and 1d are constructed analogously.

A simple look into the four plots suggests that there is a good linear correlation between I_{χ^2} and CR, GCI, CI^* respectively, whereas a weaker linear correlation is evidenced between I_{χ^2} and GW. In order to precisely evaluate this correlation, we computed Pearson's correlation coefficient ρ between I_{χ^2} and each one of the remaining indices. We used the set $S = \{\mathbf{B}_1, ..., \mathbf{B}_N\}$ of PCMs described in Subsection 5.1, with $N = 10,000$. Given two inconsistency indices, say I_i and I_j, let us consider the two associated sets of inconsistency values $\{I_i(\mathbf{B}_1), ..., I_i(\mathbf{B}_N)\}$ and $\{I_j(\mathbf{B}_1), ..., I_j(\mathbf{B}_N)\}$. We computed the corresponding Pearson Correlation Coefficient $\rho(i,j)$ for different values of n and σ,

$$\rho(i,j) = \frac{\sum_{p=1}^{N}(I_i(\mathbf{B}_p) - \bar{I}_i)(I_j(\mathbf{B}_p) - \bar{I}_j)}{(N-1)s_i s_j} \; , \tag{15}$$

where $\bar{I}_i$ and s_i are the mean and the standard deviation of $I_i(\mathbf{B}_p)$, respectively. Analogously, $\bar{I}_j$ and s_j are the mean and the standard deviation of $I_j(\mathbf{B}_p)$. As an example, the results for $n = 6$ and $\sigma = 0.4$ are reported in Table 5.
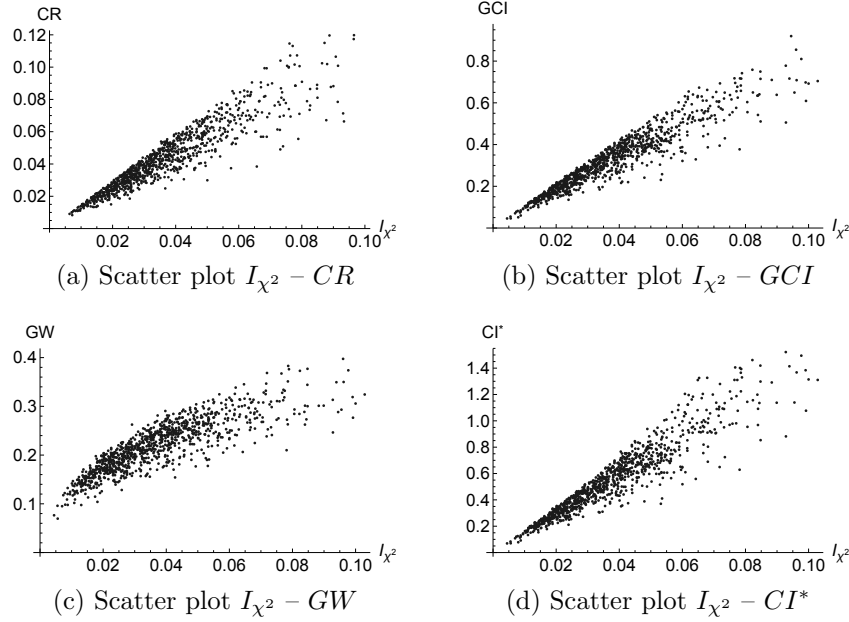
Figure 1: Graphical comparison between I_{χ^2} and four other indices, case 1.

ρ	CR	CI^*	GW	GCI
I_{χ^2}	0.9264	0.9238	0.8333	0.9255

Table 5: Pearson's correlation coefficient ρ between I_{χ^2} and the other four indices, case 1.

Clearly, the numerical values in Table 5 confirm the previous remarks on the linear correlations between the indices.

Pearson's correlation coefficient ρ measures the *linear* correlation between pairs of observations. Nevertheless, two sequences of data may be strongly related even if the relationship is highly non linear. Spearman's rank correlation coefficient, for example, measures to which extent the pairs of observations are comonotone, no matter the linearity of the relationship. More precisely, Spearman coefficient compares the way two indices rank the matrices of a fixed set of PCMs. If two indices produce the same ranking, from the best matrix to the worst one, then Spearman coefficient reaches its maximum value 1. This means that the two indices are related by a monotone increasing function. Let $r(I_i(\mathbf{B}_p))$ be the rank of matrix $\mathbf{B}_p$ according to inconsistency index I_i and $r(I_j(\mathbf{B}_p))$ be the rank of matrix $\mathbf{B}_p$ according to

inconsistency index I_j. Note that each index I_i (or I_j) induces a ranking, or ordering, in the set S. For example, $r(I_i(\mathbf{B}_p)) = 1$ means that index I_i put matrix $\mathbf{B}_p$ at the first place, i.e. it considers $\mathbf{B}_p$ the most consistent one in the set. Spearman coefficient is defined as

$$\varrho(i,j) = 1 - \frac{6\sum_{p=1}^{N} d_p^2}{N(N^2 - 1)}, \tag{16}$$

where $d_p^2 = [r(I_i(\mathbf{B}_p)) - r(I_j(\mathbf{B}_p))]^2$. For the set S of PCMs described above, we computed Spearman coefficient (16) between I_{χ^2} and the other four indices. As an example, we again report the results for $n = 6$ and $\sigma = 0.4$ in Table 6.

ϱ	CR	CI^*	GW	GCI
I_{χ^2}	0.9473	0.9469	0.8659	0.9470

Table 6: Spearman's coefficient ϱ between I_{χ^2} and the other four indices, case 1.

From the data in Table 6, it can be observed that also for what concerns comonotonicity, there is a good agreement between I_{χ^2} and CR, CI^*, GCI, respectively. Similarly to what has been noted for linear correlation, there is a weaker relationship between I_{χ^2} and GW.

Case 2: Randomly generated Pairwise Comparison Matrices

We repeated the same study as in case 1, but with a set of randomly generated reciprocal PCMs. More precisely, we generated a set $T = \{\mathbf{A}_1, ..., \mathbf{A}_N\}$ with $N = 10,000$ PCMs of order 6 by randomly sampling the upper diagonal entries from Saaty's scale

$$\{1/9, 1/8, \ldots, 1/2, 1, 2, \ldots, 8, 9\},$$

and consequently we computed the lower diagonal entries according to reciprocity, $a_{ji} = 1/a_{ij}$. On this matrix set T we performed the graphical comparisons by means of scatter plots, we computed Pearson's correlation coefficients and Spearman's rank correlation coefficients similarly to what we did in case 1. The results are reported in Figure 2, Table 7 and Table 8.

By comparing Figure 1 and Figure 2, we can observe that the numerical values in the latter are obviously much larger than the ones in the former. Nevertheless, the reported shapes are sufficiently similar to confirm the type
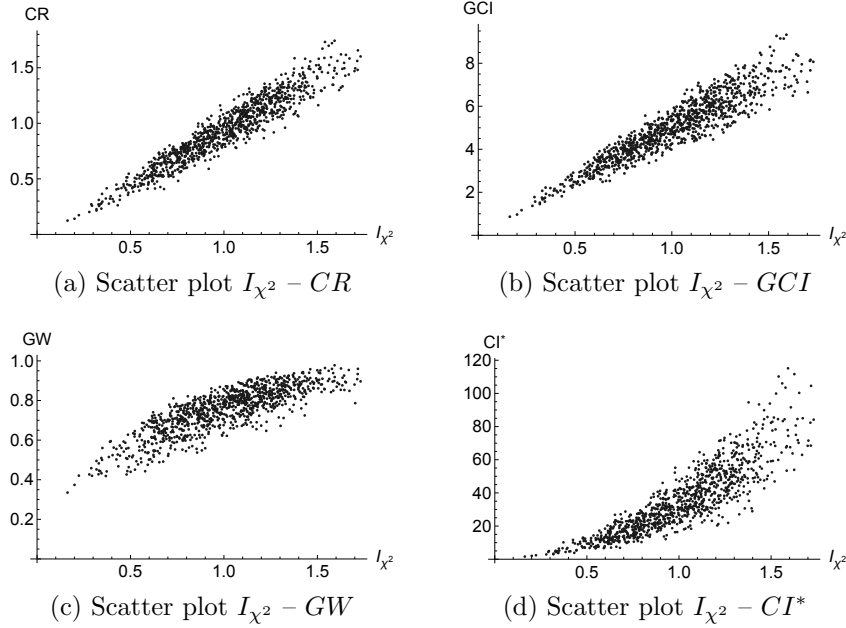
17

(a) Scatter plot $I_{\chi^2} - CR$ (b) Scatter plot $I_{\chi^2} - GCI$

(c) Scatter plot $I_{\chi^2} - GW$ (d) Scatter plot $I_{\chi^2} - CI^*$

Figure 2: Graphical comparison between I_{χ^2} and four other indices, case 2.

ρ	CR	CI^*	GW	GCI
I_{χ^2}	0.9534	0.8764	0.8522	0.9348

Table 7: Pearson's correlation coefficient ρ between I_{χ^2} and the other four indices, case 2.

of relationship between indices even for different sets of PCMs. Note that in Figure 2 the mean values of I_{χ^2} and CR are both equal to 1. This is coherent with definitions (5) and (9) where the indices are normalized with their random value. The results in Tables 5, 6, 7 and 8 confirm a strong relationship between I_{χ^2} and CR. The relationship between I_{χ^2} and GCI is slightly weaker but still good. The same holds for CI^*, whereas the weakest relationship is between I_{χ^2} and GW.

6   Final remarks and future work

The interest of the scientific community on the problem of consistency evaluation is evidenced by the large number of papers recently published on this topic in highly rated journals. Our proposal of an alternative way of measur-

ϱ	CR	CI^*	GW	GCI
I_{χ^2}	0.9533	0.9054	0.8529	0.9362

Table 8: Spearman's coefficient ϱ between I_{χ^2} and the other four indices, case 2.

ing inconsistency through the newly introduced index I_{χ^2} is not focused on proving that the already known indices behave worse than ours. The aim of our paper is to contribute to a more deep understanding of some important aspects of the problem, being aware that all the proposed indices have pros and cons. We consider as particularly interesting the fact that inconsistency indices defined with completely different approaches share very similar properties. As an example, we verified this fact by considering I_{χ^2} and Saaty's CR. Nevertheless, this interesting outcome apply also to the index studied in (Grzybowski, 2016). In our view, this coherent behavior supports from independent points of view what can be considered a common way of measuring inconsistency. Future research could clarify this point. The proof that I_{χ^2} satisfies properties (P3) and (P4) still remains an open problem that we will investigate in future. In our future research work, we also aim to study the computational complexity of index (9) and to compare it with that of other indices. We plan to extend the comparison study described in Section 5.2 to other inconsistency indices too, in order to have more complete and reliable outcomes.

References

Aguarón J and Moreno-Jiménez J M (2003). The geometric consistency index: approximated thresholds. European Journal of Operational Research 147(1): 137–145.

Alonso J A and Lamata M T (2006). Consistency in the Analytic Hierarchy Process: a new approach. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 14(4): 445–459.

Bana e Costa C A and Vansnick J-C (2008). A critical analysis of the eigenvalue method used to derive priorities in AHP. European Journal of Operational Research 187(3): 1422–1428.

Barzilai J (1998). Consistency measures for pairwise comparison matrices. Journal of Multi-Criteria Decision Analysis 7(3): 123–132.

Barzilai J (2005). Measurement and preference function modelling. International Transactions in Operational Research 12(2): 173–183.

Bozóki S and Rapcsák T (2008). On Saatys and Koczkodajs inconsistencies of pairwise comparison matrices. Journal of Global Optimization 42(2): 157–175.

Brunelli M (2017). Studying a set of properties of inconsistency indices for pairwise comparisons. Annals of Operations Research 248(1): 143-161.

Brunelli M and Fedrizzi M (2015). Axiomatic properties of inconsistency indices for pairwise comparisons. Journal of the Operational Research Society 66(1): 1–15.

Brunelli M, Canal L and Fedrizzi M (2013). Inconsistency indices for pairwise comparison matrices: a numerical study. Annals of Operations Research 211(1): 493–509.

Brunelli M, Critch A and Fedrizzi M (2013). A note on the proportionality between some consistency indices in the AHP Applied Mathematics and Computation 219(14): 7901–7906.

Cavallo B and D'Apuzzo L (2009). A general unified framework for pairwise comparison matrices in multicriterial methods. International Journal of Intelligent Systems 24(4): 377–398.

Cavallo B, D'Apuzzo L and Squillante M (2012). About a consistency index for pairwise comparison matrices over a divisible alo-group. International Journal of Intelligent Systems 27(2): 153–175.

Choo E U and Wedley W C (2004). A common framework for deriving preference values from pairwise comparison matrices. Computers and Operations Research 31(6): 893–908.

Crawford G and Williams C (1985). A note on the analysis of subjective judgement matrices. Journal of Mathematical Psychology 29(4): 387–405.

Csató L (2017). Characterization of an inconsistency ranking for pairwise comparison matrices. Annals of Operations Research, published online: August 2017, DOI 10.1007/s10479-017-2627-8.

Dijkstra T K (2013). On the extraction of weights from pairwise comparison matrices. Central European Journal of Operations Research 21(1): 103–123.

Duszak Z and Koczkodaj W W (1994). Generalization of a new definition of consistency for pairwise comparisons. Information Processing Letters 52(5): 273–276.

Fedrizzi M and Brunelli M (2009). Fair consistency evaluation for reciprocal relations and group decision making. New Mathematics and Natural Computation 5(2): 407–420.

Fedrizzi M, Fedrizzi M and Marques Pereira R A (2002). On the issue of consistency in dynamical consensual aggregation. In: Bouchon Meunier B, Gutierrez Rios J, Magdalena L and Yager R R (eds). Technologies for Constructing Intelligent Systems 1, Studies in Fuzziness and Soft Computing 89, 2002 Springer, Physica, Heidelberg, pp. 129–137.

Golden B L and Wang Q (1989). An alternate measure of consistency. In: Golden B L, Wasil E A and Harker P T (eds). The Analytic Hierarchy Process, Applications and studies. Springer–Verlag, Berlin–Heidelberg, pp. 68–81.

Grzybowski A Z (2012). Note on a new optimization based approach for estimating priority weights and related consistency index. Expert Systems With Applications 39(14): 11699–11708.

Grzybowski A Z (2016). New results on inconsistency indices and their relationship with the quality of priority vector estimation. Expert Systems With Applications 43: 197–212.

Jensen R E (1983). Comparison of eigenvector, least squares, Chi square and logarithmic least squares methods of scaling a reciprocal matrix. Trinity University, Working Paper 127.

Kazibudzki P T (2016). An examination of performance relations among selected consistency measures for simulated pairwise judgments. Annals of Operations Research 244(2): 525-544.

Koczkodaj W W (1993). A new definition of consistency of pairwise comparisons. Mathematical & Computer Modelling 18(7): 79–84.

Koczkodaj W W and Szwarc R (2014). On axiomatization of inconsistency indicators for pairwise comparisons. Fundamenta Informaticae 132(4): 485–500.

Kou G and Lin C (2014). A cosine maximization method for the priority vector derivation in AHP. European Journal of Operational Research 235(1): 225–232.

Lin C C (2007). A revised framework for deriving preference values from pairwise comparison matrices. European Journal of Operational Research 176(2): 1145–1150.

Lin C, Kou G and Ergu D (2013). An improved statistical approach for consistency test in AHP. Annals of Operations Research 211(1): 289–299.

Lipovetsky S and Conklin W M (2002). Robust estimation of priorities on the AHP. European Journal of Operational Research 137(1): 110–122.

Mazurek J (2017). On inconsistency indices and inconsistency axioms in pairwise comparisons. ArXiv:1703.05204v2.

Pearson K (1920). Notes on the history of correlation. Biometrika 13(1): 25–45.

Peláez J I and Lamata M T (2003). A new measure of inconsistency for positive reciprocal matrices. Computer and Mathematics with Applications 46(12): 1839–1845.

Ramík J and Korviny P (2010). Inconsistency of pair-wise comparison matrix with fuzzy elements based on geometric mean. Fuzzy Sets and Systems 161(11): 1604–1613.

Saaty T L (1977). A scaling method for priorities in hierarchical structures. Journal of Mathematical Psychology 15(3): 234–281.

Saaty T L (1994a). Highlights and critical points in the theory and application of the Analytic Hierarchy Process. European Journal of Operational Research 74(3): 426–447.

Saaty T L (1994b). Fundamentals of Decision making and priority theory with the AHP. RWS publications, Pittsburg.

Salo A and Hämäläinen R (1995). Preference programming through approximate ratio comparisons. European Journal of Operational Research 82(3): 458–475.

Shiraishi S, Obata T and Daigo M (1998). Properties of a positive reciprocal matrix and their application to AHP. Journal of the Operations Resesearch Society of Japan 41(3): 404–414.

Snedecor G W and Cochran W G (1980). Statistical Methods. Iowa State University Press.

Spearman C E (1904). The proof and measurement of association between two things. The American Journal of Psychology 15(1): 72–101.

Stein W E and Mizzi P J (2007). The harmonic consistency index for the analytic hierarchy process. European Journal of Operational Research 177(1): 488–497.

Temesi J (2011). Pairwise comparison matrices and error-free property of the decision maker. Central European Journal of Operational Research 19(2): 239–249.

Wang Y-M, Fan Z-P and Hua Z (2007). A chi-square method for obtaining a priority vector from multiplicative and fuzzy preference relations. European Journal of Operational Research 182(1): 356–366.

Wu Z and Xu J (2012). A consistency and consensus based decision support model for group decision making with multiplicative preference relations. Decision Support Systems 52(3): 757-767.

Xu Y, Chen L, Li K W and Wang H (2015). A chi-square method for priority derivation in group decision making with incomplete reciprocal preference relations. Information Sciences 306: 166–179.

Xu Z S (2000). Generalized Chi-square method for the estimation of weights. Journal of Optimization Theory and Applications 107(1): 183–192.