Ann Oper Res (2013) 211:493–509
DOI 10.1007/s10479-013-1329-0

Inconsistency indices for pairwise comparison matrices: a numerical study

Matteo Brunelli · Luisa Canal · Michele Fedrizzi

Published online: 27 February 2013

Abstract Evaluating the level of inconsistency of pairwise comparisons is often a crucial step in multi criteria decision analysis. Several inconsistency indices have been proposed in the literature to estimate the deviation of expert's judgments from a situation of full consistency. This paper surveys and analyzes ten indices from the numerical point of view. Specifically, we investigate degrees of agreement between them to check how similar they are. Results show a wide range of behaviors, ranging from very strong to very weak degrees of agreement.

Keywords Pairwise comparison matrices · Consistency · Analytic hierarchy process

1 Introduction

In the field of studies on the human behavior and decision processes, a large amount of papers investigate consistency in judgments and evaluations (for a recent review, see Karelaia and Hogarth 2008). The measurement of the quality of expertise of the decision makers has been a matter of much debate, especially in domains lacking external criteria which enable verification (for a discussion, see Shanteau et al. 2003). Consensus exists

M. Brunelli (✉)
Systems Analysis Laboratory, Department of Mathematics and Systems Analysis, Aalto University, Espoo, Finland
e-mail: matteo.brunelli@aalto.fi

M. Brunelli
Institute for Advanced Management Systems Research, Åbo Akademi University, Åbo, Finland

L. Canal
Department of Cognitive Sciences, University of Trento, Rovereto, Italy
e-mail: luisa.canal@unitn.it

M. Fedrizzi
Department of Industrial Engineering, University of Trento, Trento, Italy
e-mail: michele.fedrizzi@unitn.it

among some traits of experts' judgments including selectivity in information search/use (Phelps and Shanteau 1978; Shanteau 1989; Slovic 1969), discrimination among different stimuli (Weiss and Shanteau 2003), and intra-individual consistency (Weiss et al. 2006; Karelaia and Hogarth 2008). Unfortunately, there is evidence that many professional judgments do not fulfill these requirements and consequently their stated judgments are not consistent.

In decision making processes, when an expert has to grade alternatives subjectively, in order to form a rating, he/she is often asked to express his/her preferences by means of cardinal preference relations. In this framework, the decision maker is then asked to pairwise compare alternatives and, for each pair, to express his degree of preference of the first alternative over the second, or vice versa. This approach was formally introduced by Thurstone (1927) and its importance was highlighted by Koczkodaj (1993). Although there is not a unique way for representing subjective preferences under the form of pairwise comparisons over pairs, it is possible to define a consistency condition such that, if it holds, the decision maker is considered coherent and his/her judgments are not contradictory. Namely, a decision maker is consistent if his/her opinions respect some cardinal transitivity conditions of preferences on triplets of alternatives. Nevertheless, it is well known that in making paired comparisons, people do not have the intrinsic logical ability to always be consistent (Saaty 1994). Despite some contrary opinions, e.g. Linares (2009), consistency has always been regarded as a desirable property. Consequently, for each type of representation of preferences, there is a meeting of minds on the right definition of consistency, whereas there is a rather large number of proposals to evaluate the amount of inconsistency contained in the preferences. The firstly introduced and most popular index for measuring inconsistency is the Consistency Index, defined by Saaty (1977, 1980) in the Analytic Hierarchy Process (AHP). Originally introduced as a tool for estimating degrees of membership in fuzzy sets (Saaty 1974), the AHP is a method for decisions which draws heavily on the theory of preference relations. Within the framework of the AHP, consistency evaluation has probably been one of the most debated issues. Since there is no consensually accepted method for measuring inconsistency, every index proposed in the literature is itself a different definition of inconsistency degree. Since there is not a satisfactory comparative study in the literature, in this paper we numerically investigate the agreement between different indices in evaluating inconsistency of cardinal preference relations.

Note that the relevance of inconsistency indices goes beyond the quantification of inconsistency as they are also often used as methods to complete pairwise comparison matrices with missing elements.

This paper is outlined as follows. In Sect. 2 we introduce pairwise comparison matrices and the concepts underlying the idea of inconsistency, meant as a deviation from a situation of perfect consistency. In Sect. 3 we recall the definitions of the ten inconsistency indices that we are going to analyze numerically. Section 4 describes the numerical study and exposes the results. In Sect. 5, we comment and discuss the findings of Sect. 4.

2 Pairwise comparison matrices and their inconsistency

Given a non-empty finite set of alternatives $X = \{x_1, \dots, x_n\}$, a pairwise comparison matrix (*PCM*) is a positive real valued matrix $\mathbf{A} = (a_{ij})_{n\times n}$ with (i) $a_{ii} = 1\ \forall i$ and (ii) $a_{ij}a_{ji} = 1$ $\forall i, j$. Despite this definition, in his papers on the AHP, Saaty (1977) claimed that, for psychological and behavioral reasons (Miller 1956), decision makers should use the bounded discrete scale $1/9, 1/8, \dots, 1/2, 1, 2, \dots, 8, 9$ instead of the set of all positive real numbers.

We are going to use this restrictive approach also because it allows us to generate random matrices. A *PCM* is said to be fully consistent if and only if the following transitivity condition holds

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

Moreover, if and only if $\mathbf{A}$ is consistent, then there exists a vector $\mathbf{w} = (w_1, \dots, w_n)$ such that

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

If $\mathbf{A}$ is consistent, then vector $\mathbf{w}$ can be obtained by using the geometric mean method

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

It is clear that a definition of consistency has been given. However, the same cannot be said for inconsistency, which is roughly seen as a deviation from the condition of full consistency. A number of indices have been proposed in the literature to quantify the extent of this deviation and they will be presented in the next section. Let us just recall that, in general, an inconsistency index is a function

$$I : \mathcal{A} \to \mathbb{R}$$

where $\mathcal{A}$ is the set of all the *PCM*s, i.e. $\mathcal{A} = \{\mathbf{A} = (a_{ij})_{n\times n} | a_{ij} > 0,\ a_{ij}a_{ji} = 1\ \forall i, j,\ n > 2\}$. In words, an inconsistency index is a function which associates *PCM*s with real numbers, where the real numbers are quantifications of inconsistency. Consequently, different indices are not alternative ways for evaluating a well-defined quantity. Instead, different indices represent different definitions of the degree of inconsistency of pairwise comparison matrices $\mathbf{A} \in \mathcal{A}$.

3 Inconsistency indices

In this study we consider ten inconsistency indices, but there exist alternative approaches. For instance, Salo (1993) and Salo and Hämäläinen (1995) considered that ambiguity of judgments goes arm-in-arm with their inconsistency and therefore introduced, and then pushed forward, an ambiguity index which can also be used as an estimation of inconsistency. Osei-Bryson (2006) proposed an interpretable parametric optimization problem where the value of the objective function can be used as an inconsistency index. Unlike other inconsistency indices, the index proposed by Osei-Bryson (2006) allows the decision maker to establish a priori some acceptability thresholds for the inconsistencies associated with single entries of the *PCM*.

Although we shall keep the description of the indices self-contained—as their full illustration would be beyond the scope of this paper—we skip the details and invite the reader to refer to the original works. We only preliminarily note that some indices were originally called consistency indices whereas some others were called inconsistency indices. Here we maintain the original names but we clarify that, in spite of these terminological difficulties, the nature of the indices is the same, i.e. the greater the value of the index, the greater the inconsistency of the *PCM*.

3.1 CI and CR

According to the result that given a *PCM* **A**, its maximum eigenvalue, $\lambda_{\max}$, is equal to n if and only if the matrix is consistent (and greater than n otherwise), Saaty (1977) proposed a consistency index

$$CI = \frac{\lambda_{\max} - n}{n-1}. \tag{4}$$

However, empirical analyses confirmed that the expected value of *CI* of a random matrix of size $n+1$ is, in average, greater than the expected value of *CI* of a random matrix of order n. Consequently, *CI* is not reliable in comparing matrices of different size. Therefore, it needs to be rescaled.

CR, which stands for *Consistency Ratio*, is the standardized version of *CI*. Given a matrix of order n, *CR* can be obtained dividing *CI* by a real number *RI* (*Random Index*) which is nothing else but the average *CI* obtained from a large enough set of randomly generated matrices of size n. Hence,

$$CR = \frac{CI}{RI}. \tag{5}$$

3.2 Index of determinants and c_3

This index (Pelàez and Lamata 2003) is mainly based on the following property of a *PCM* of order three. Expanding the determinant of a 3×3 real matrix one obtains

$$\det(\mathbf{A}) = \frac{a_{ik}}{a_{ij}a_{jk}} + \frac{a_{ij}a_{jk}}{a_{ik}} - 2.$$

If the matrix is not consistent, then $\det(\mathbf{A}) > 0$, because $\frac{a}{b} + \frac{b}{a} - 2 > 0 \ \forall a \neq b, a, b > 0$.

It is possible to generalize this result for matrices of order greater than three and define an inconsistency index as the arithmetic mean of the determinants of all the possible submatrices $\mathbf{T}_{ijk}$ of a given *PCM*, constructed in a way so that they respect the following formulation

$$\mathbf{T}_{ijk} = \begin{pmatrix} 1 & a_{ij} & a_{ik} \\ a_{ji} & 1 & a_{jk} \\ a_{ki} & a_{kj} & 1 \end{pmatrix}, \quad \forall i < j < k.$$

The number of so constructed submatrices is $\binom{n}{3} = \frac{n!}{3!(n-3)!}$. The result is a standardized index and its value is the average inconsistency computed for all the submatrices $\mathbf{T}_{ijk}$ $(i < j < k)$

$$CI^* = \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} \left(\frac{a_{ik}}{a_{ij}a_{jk}} + \frac{a_{ij}a_{jk}}{a_{ik}} - 2 \right) \Big/ \binom{n}{3}. \tag{6}$$

The coefficient c_3 of the characteristic polynomial of a *PCM* was also proposed to act as an inconsistency index by Shiraishi and Obata (2002) and Shiraishi et al. (1998, 1999). In fact, by definition, the characteristic polynomial has the following form

$$P_{\mathbf{A}}(\lambda) = \lambda^n + c_1\lambda^{n-1} + \cdots + c_{n-1}\lambda + c_n,$$

with $c_1, \ldots, c_n$ that are real numbers and λ the unknown. Shiraishi et al. (1998) proved that, if $c_3 < 0$, then the matrix at issue cannot be fully consistent. In fact, this is evident if one reckons that—in light of the Perron-Frobenius theorem—the only possible formulation of the characteristic polynomial which yields $\lambda_{\max} = n$, is

$$P_{\mathbf{A}}(\lambda) = \lambda^{n-1}(\lambda - n). \tag{7}$$

Thus, the presence of c_3 is certainly a symptom of inconsistency. Moreover, they also proved that c_3 has the following analytic expression

$$c_3 = \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} \left(2 - \frac{a_{ik}}{a_{ij}a_{jk}} - \frac{a_{ij}a_{jk}}{a_{ik}}\right). \tag{8}$$

Since it was showed that indices CI^* (6) and c_3 (8) are proportional (Brunelli et al. 2013), then they are equivalent, and, to avoid redundancy, in the following we will consider them as a unique index.

3.3 Squared differences index

The definition of this index (Chu et al. 1979) is based on characterization (2) and it assumes that each deviation from the desirable situation should be considered a symptom of inconsistency. Thus, the sum of the squares of the deviations $(a_{ij} - w_i/w_j)$ $\forall i \neq j$ is here considered a fair and global quantification of inconsistency

$$LS = \min_{w_1,\ldots,w_n} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} \left(a_{ij} - \frac{w_i}{w_j}\right)^2 \quad \text{s.t.} \quad \sum_{i=1}^{n} w_i = 1, \quad w_i > 0. \tag{9}$$

Index LS, which stands for least squares, is also easy to be standardized since the number of non-diagonal terms in the sum, as noted above, is $n(n-1)$. Let us note that the argument minimizing (9) is the priority vector $\mathbf{w}^* = (w_1^*, \ldots, w_n^*)$ associated with the pairwise comparison matrix $\mathbf{W}^* = (w_i^*/w_j^*)_{n\times n}$ which minimizes the Frobenius norm $\|\mathbf{A} - \mathbf{W}\|_2$ with $\mathbf{W} = (w_i/w_j)_{n\times n}$. Despite the elegant formulation, optimization problem (9) is tough to be solved numerically, multiple solutions can exist and no exact method or analytic solution has been found to solve it. Bózoki (2008) built an equation system whose roots yield to the optimal components of $\mathbf{w}$. However, even this method, suffers of a huge computational complexity. In order to overcome this problem, Anholcer et al. (2011) proposed some simplifications which are based on some uncertain assumptions.

3.4 Geometric consistency index

This index had been implicitly introduced by Crawford and Williams (1985), then it was reexamined by Aguaròn and Moreno-Jimènez (2003). It considers the priority vector to be estimated by means of the geometric mean method (3). With the estimated weights it is possible to build a local estimator of inconsistency,

$$e_{ij} = a_{ij}\frac{w_j}{w_i}, \quad i, j = 1, \ldots, n. \tag{10}$$

For consistent matrices the value of e_{ij} is equal to 1 because it is the result of a multiplication of an entry times its reciprocal. Therefore, since $a_{ij} = \frac{w_i}{w_j} \Rightarrow \ln e_{ij} = 0$, it is then possible to define a global inconsistency index, i.e. the Geometric Consistency Index (GCI), that is

$$GCI = \frac{2}{(n-1)(n-2)} \sum_{i=1}^{n} \sum_{j=i+1}^{n} \ln^2 e_{ij}. \tag{11}$$

Furthermore, Brunelli et al. (2013) proved that GCI is linearly proportional to the following quantity

$$\rho = \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} (\log_9 a_{ik} a_{kj} a_{ji})^2, \tag{12}$$

which means that such index can be expressed as a measure of deviation from either (1) or (2).

3.5 Harmonic consistency index

If and only if $\mathbf{A}$ is a consistent *PCM*, then its columns are proportional and $\text{rank}(\mathbf{A}) = 1$. Therefore, it is fair to suppose that the less proportional are the columns, the less consistent is the matrix. An index of inconsistency based on proportionality between columns was then proposed by Stein and Mizzi (2007) and based on this sufficient condition of consistency. Given a matrix $\mathbf{A}$, the authors proposed to construct an auxiliary vector $\mathbf{s} = (s_1, \ldots, s_n)$ with $s_j = \sum_{i=1}^n a_{ij}\ \forall j$. It was proven that $\sum_{j=1}^n s_j^{-1} = 1$ if and only if $\mathbf{A}$ is consistent, and smaller than 1 otherwise. The harmonic mean of the components of vector $\mathbf{s}$ is then the result of the following

$$HM = \frac{n}{\sum_{j=1}^n \frac{1}{s_j}}. \tag{13}$$

HM itself could be an index of inconsistency, but the authors, according to computational experiments, proposed a normalization in order to align its behavior with that of *CI*. The Harmonic Consistency Index is then

$$HCI = \frac{(HM - n)(n+1)}{n(n-1)}. \tag{14}$$

3.6 Cavallo-D'Apuzzo

Besides proposing a general framework based on abelian linearly ordered groups for some representations of cardinal preferences Cavallo and D'Apuzzo (2009, 2010) introduced an approach based on some new metrics and the following inconsistency index

$$I_{CD} = \prod_{i=1}^{n-2} \prod_{j=i+1}^{n-1} \prod_{k=j+1}^{n} \left(\max\left\{ \frac{a_{ik}}{a_{ij}a_{jk}}, \frac{a_{ij}a_{jk}}{a_{ik}} \right\} \right)^{\frac{1}{\binom{n}{3}}}. \tag{15}$$

Remarkably, thanks to the fact that the justification of this index has been grounded on group theory, the authors equivalently formulated it for other well-known types of preference relations.

3.7 Relative error

The relative error index was formulated by Barzilai (1998) and requires the construction of an auxiliary matrix $\mathbf{A}^+ = (a_{ij}^+)_{n\times n} = (\log_2 a_{ij})_{n\times n}$ which is skew symmetric and represents an 'additive' *PCM* and to derive a weight vector $\mathbf{w}^+ = (w_1^+, \ldots, w_n^+)$ with $w_i^+ = \frac{1}{n}\sum_{j=1}^n a_{ij}^+$. Having done this, the consistent part of $\mathbf{A}^+$ is obtained as $\mathbf{C} = (c_{ij})_{n\times n} = (w_i^+ - w_j^+)_{n\times n}$. Another matrix, $\mathbf{E} = (e_{ij})_{n\times n} = (a_{ij} - c_{ij})_{n\times n}$, is also obtained to represent the inconsistent part of $\mathbf{A}$, such that $\mathbf{C} + \mathbf{E} = \mathbf{A}$. At this point the relative error index is derived as

$$RE = \frac{\sum_{i=1}^n \sum_{j=1}^n e_{ij}^2}{\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2}. \tag{16}$$

3.8 Koczkodaj index

Koczkodaj (1993) proposed to estimate the inconsistency of a *PCM* of order three as

$$K_{ijk} = \min\left\{\frac{1}{a_{ij}}\left|a_{ij} - \frac{a_{ik}}{a_{jk}}\right|, \frac{1}{a_{ik}}|a_{ik} - a_{ij}a_{jk}|, \frac{1}{a_{jk}}\left|a_{jk} - \frac{a_{ik}}{a_{ij}}\right|\right\}. \tag{17}$$

This method was then generalized by Duszak and Koczkodaj (1994) for $n \geq 3$,

$$K = \max\{K_{ijk} | 1 \leq i < j < k \leq n\}. \tag{18}$$

Note that, for sake of simplicity, it was proven (Duszak and Koczkodaj 1994) that (17) collapses into the following

$$K_{ijk} = \min\left\{\left|1 - \frac{a_{ik}}{a_{ij}a_{jk}}\right|, \left|1 - \frac{a_{ij}a_{jk}}{a_{ik}}\right|\right\}. \tag{19}$$

3.9 Golden-Wang index

Golden and Wang (1989) assumed that the priority vector be computed thanks to the row geometric mean and then normalized such that its components sum up to one. They call the so obtained vector $\mathbf{g}^* = (g_1^*, \dots, g_n^*)$. The same operation is similarly repeated on the columns of $\mathbf{A}$. Namely, entries on the jth column are divided by a constant $k_j = \sum_{i=1}^n a_{ij}$ and the new matrix can then be called $\mathbf{A}^* = (a_{ij}^*)$. At this point, the index is as follows,

$$GW = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\left|a_{ij}^* - g_i^*\right|. \tag{20}$$

3.10 Ramík-Korviny index

Ramík and Korviny (2010) and Ramík and Perzina (2010) presented an inconsistency index for *PCM*s whose entries are triangular fuzzy numbers. However, as they treat it as a more general case, their index can be adapted to work with *PCM*s with real valued entries. They proposed to use the geometric mean method to estimate the priority vector (3) and formulated the inconsistency index as follows,

$$NI_n^\sigma = \gamma_n^\sigma \max_{i,j}\left\{\left|a_{ij} - \frac{w_i}{w_j}\right|\right\}, \tag{21}$$

where

$$\gamma_n^\sigma = \begin{cases} \frac{1}{\max\{\sigma - \sigma^{\frac{2-2n}{n}}, \sigma^2((\frac{2}{n})^{\frac{2}{n-2}} - (\frac{2}{n})^{\frac{n}{n-2}})\}}, & \text{if } \sigma < (\frac{n}{2})^{\frac{n}{n-2}}, \\ \frac{1}{\max\{\sigma - \sigma^{\frac{2-2n}{n}}, \sigma^{\frac{2n-2}{n}} - \sigma\}}, & \text{if } \sigma \geq (\frac{n}{2})^{\frac{n}{n-2}}, \end{cases}$$

is a positive normalization factor. Let us incidentally note that this index has been criticized by Brunelli (2011) and an alternative metric was proposed.

4 Numerical study on indices' agreement

The main aim of this paper is to compare different inconsistency evaluation proposals. All the ten different indices described in the previous section aim to measure exactly the same notion, i.e. the deviation of a given *PCM* from consistency. As a consequence, some questions naturally arise: is there a good agreement between different indices? Do they classify pairwise comparison matrices in a similar way? Is their use interchangeable? Can the choice of different inconsistency indices affect the result of a decision process? In our opinion, it is important to give reliable answers to these questions, since many decision models proposed in the literature assume that consistency plays an important role in the decision process. In group decision making, for instance, the aggregation of the individual preferences can be performed taking into account their closeness to consistency. In this section we study the inconsistency indices from a statistical perspective through numerical simulations. We are aware of the difficulties in comparing indices defined in different frameworks. In particular, the use of bounded/unbounded scales leads to different consequences on consistency. As specified in Sect. 2, in our study we chose to use the Saaty's bounded scale due to its importance and popularity. By modifying, as described below, initial values of this scale, it is clearly possible to obtain different real numbers. We performed our study on *PCM*s of order 4, 6 and 8 since these dimensions fits well real world applications. For reasons of space, we report in this section only the results corresponding to *PCM*s of order 6. Then, in Sect. 5, we briefly discuss and compare the results obtained with *PCM*s of different order.

First, in Sect. 4.1, we present the scatter plots reporting the values of the indices taken two by two. Then, we focus on three statistical tools. In Sect. 4.2, we compute the Pearson's correlation coefficient for each pair of indices in order to highlight the linear correlation between them. In Sect. 4.3, by means of the Spearman's rank correlation coefficient, we study the comonotonicity between the indices. In Sect. 4.4, by means of Cohen's kappa coefficient, we study the agreement between indices in classifying pairwise comparison matrices according to their consistency.

We use two main different classes of *PCM*s: a set S_1 of randomly generated matrices and a set S_2 of consistent matrices perturbed by a random noise, like in Choo and Wedley (2004). With this choice, we study the consistency evaluation of the most general type of matrices as well as that of matrices possibly elicited by decision makers in real life problems. The numerical data set we use in Sects. 4.1, 4.2, 4.3 and 4.4 was constructed as follows and refers to sets S_1 and S_2.

- We generated the set $S_1 = \{\mathbf{A}_1, \ldots, \mathbf{A}_N\}$ with $N = 10000$ *PCM*s 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 computed the lower diagonal entries according to reciprocity, $a_{ji} = 1/a_{ij}$.
- We generated the second set $S_2 = \{\mathbf{B}_1, \ldots, \mathbf{B}_N\}$ with $N = 10000$ *PCM*s of order 6 obtained by means of a random perturbation on consistent *PCM*s. More precisely, the following procedure was repeated 10000 times: first, a consistent *PCM* $\mathbf{B} = (b_{ij})$ is constructed by setting $(b_{ij}) = (\frac{w_i}{w_j})$, where $(w_1, \ldots, w_6)$ 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 $b_{ij} \to b_{ij}(1 + \beta)$, where β is a random variable with normal distribution, $\beta \sim N(0, \sigma)$, with $\sigma = 0.5$. The elements below the diagonal of the *PCM* are modified accordingly to preserve reciprocity, $b_{ji} = 1/b_{ij}$.

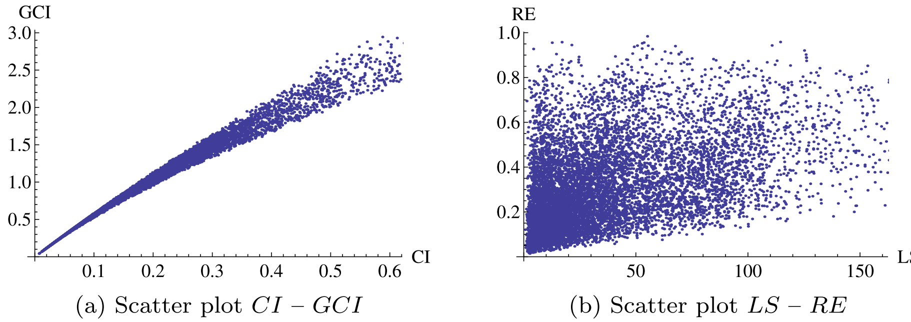

Fig. 1 Two scatter plots

– We computed the 10 inconsistency indices $I_1(\mathbf{A}_p), \ldots, I_{10}(\mathbf{A}_p)$, for every *PCM* in the random matrices set, $\mathbf{A}_p \in S_1$. Note that, for simplicity, we now denote the ten inconsistency indices by $I_1, \ldots, I_{10}$, whereas in Sect. 3 we used the original notation introduced by the various authors. The notation correspondence preserves the previous presentation order, namely, $I_1 = CI$, $I_2 = CI^*$, $I_3 = LS$, $I_4 = GCI$, $I_5 = HCI$, $I_6 = I_{CD}$, $I_7 = RE$, $I_8 = K$, $I_9 = GW$, $I_{10} = NI_n^\sigma$.
– Similarly, we considered set S_2 and for every *PCM* $\mathbf{B}_p \in S_2$, we computed the 10 inconsistency indices $I_1(\mathbf{B}_p), \ldots, I_{10}(\mathbf{B}_p)$.
– To summarize, our data set is composed of two 10000×10 tables containing the inconsistency values corresponding to the ten inconsistency indices for each set S_1 and S_2.

Finally, following Choo and Wedley (2004), we considered consistent matrices where only one entry was modified to make them inconsistent. The results corresponding to this last set of matrices are discussed in Sect. 4.5.

4.1 Graphical representations

By first considering set S_1, for each pair of indices $\{I_i, I_j\}$ we produced a scatter plot for the 10000 *PCM*s. Each one of the 10000 *PCM*s is represented by a point of the plot, the coordinates being the values of I_i and I_j respectively. The same study was performed on the matrices of set S_2. A cloud of dispersed points indicates a scarce relationship between the two indices, whereas a regular set of points closely disposed along a curve indicates a strong functional relationship. In Fig. 1(a), for example, indices $I_1 = CI$ and $I_4 = GCI$ are compared and the plot suggests a good agreement between the two indices. In Fig. 1(b), conversely, a poor agreement between indices $I_3 = LS$ and $I_7 = RE$ is evidenced. Both scatter plots refer to set S_2.

The scatter plots for all the comparisons among the 10 indices are compactly represented in Fig. 2 for the *PCM*s in set S_1 and in Fig. 3 for the *PCM*s in set S_2. In Figs. 2 and 3 the scatter plots of 500 *PCM*s are reported for a clearer graphical visualization. Note that the plots in the diagonal boxes are straight lines since identical values are involved. The plots evidence remarkably different behaviors of the indices and we will comment some relevant cases in the next section.

4.2 Pearson correlation coefficient

In order to study the linear correlation between pairs of inconsistency indices described in Sect. 3, we considered the Pearson Correlation Coefficient and proceeded as fol-

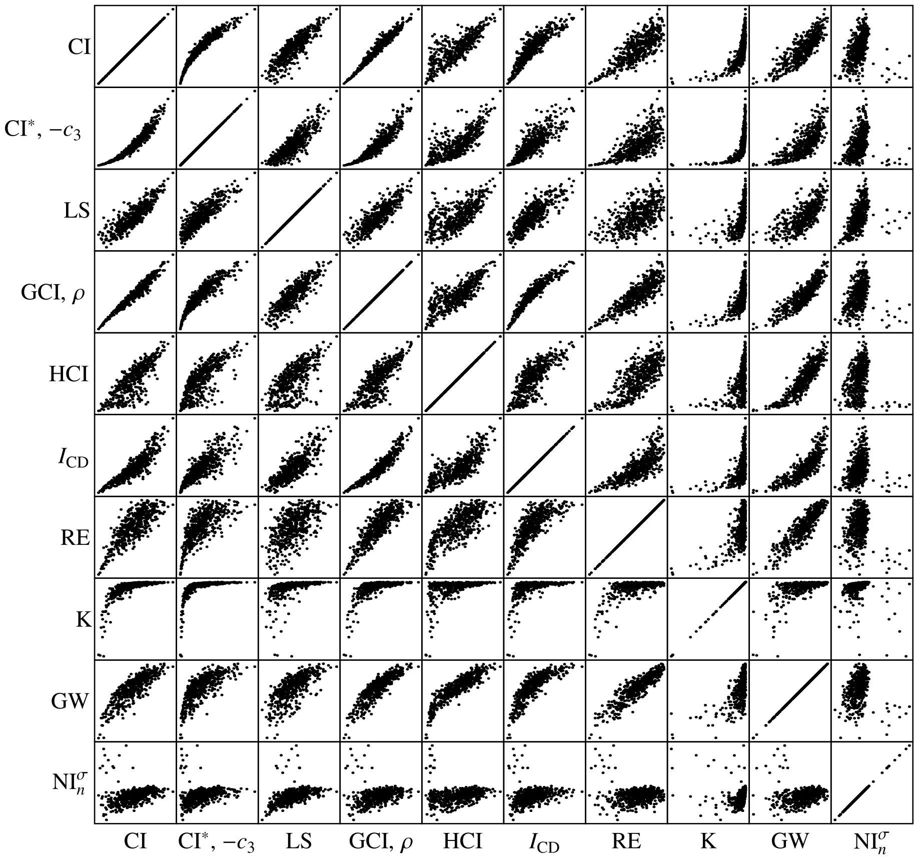

Fig. 2 Scatter plots of inconsistency indices for random matrices

lows. We first considered the set S_1 of random *PCM*s. We fixed two inconsistency indices, say I_i and I_j, and we considered the two associated sets of inconsistency values $\{I_i(\mathbf{A}_1), \dots, I_i(\mathbf{A}_N)\}$ and $\{I_j(\mathbf{A}_1), \dots, I_j(\mathbf{A}_N)\}$. We computed the corresponding Pearson Correlation Coefficient $r(i, j)$,

$$r(i, j) = \frac{\sum_{p=1}^{N} (I_i(\mathbf{A}_p) - \bar{I}_i)(I_j(\mathbf{A}_p) - \bar{I}_j)}{(N-1)s_i s_j}, \tag{22}$$

where $\bar{I}_i$ and s_i are the mean and the standard deviation of $I_i(\mathbf{A}_p)$, respectively. Analogously, $\bar{I}_j$ and s_j are the mean and the standard deviation of $I_j(\mathbf{A}_p)$. We did the same for all $\binom{10}{2}$ pairs of indices $\{I_i, I_j\}$, $i = 1, \dots, 9$; $j = i, \dots, 10$. We reported the $r(i, j)$ values in Table 1. Note that indices CI^* and $-c_3$ are reported in the same row/column due to their proportionality, as pointed out in Sect. 3. The same holds for indices *GCI* and ρ. We repeated the same study on the set S_2 and the results are reported in Table 2.

4.3 Spearman rank correlation coefficient

While the Pearson's correlation coefficient measures the *linear* correlation between pairs of observations, the Spearman index (Spearman 1904; Snedecor and Cochran 1980) measures

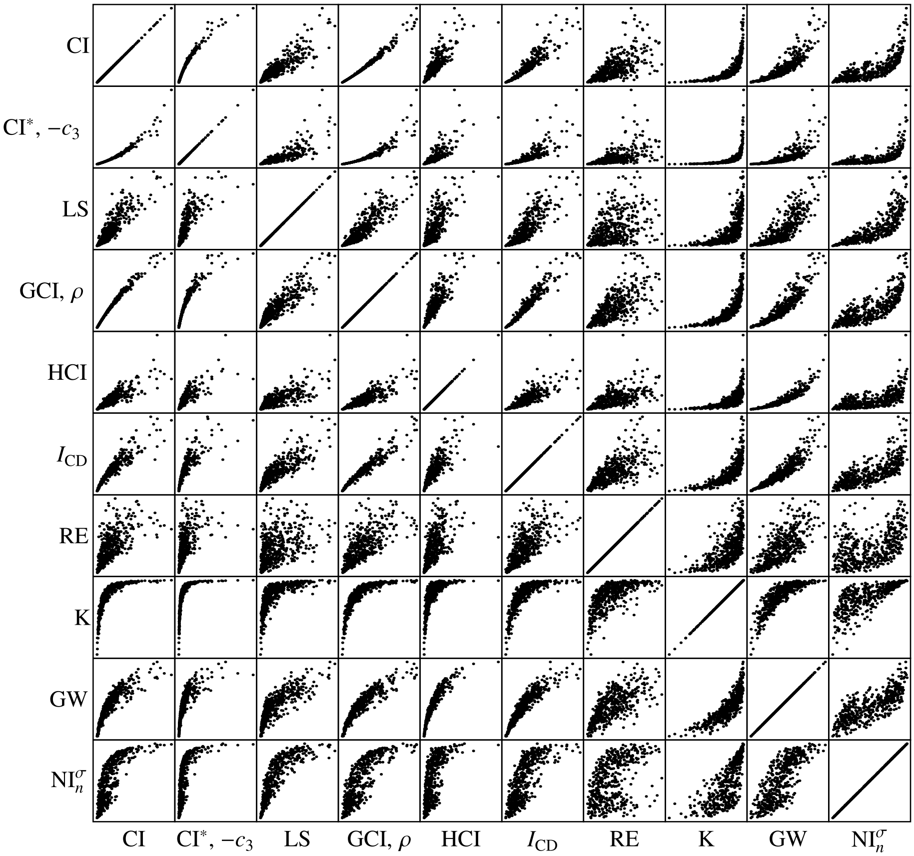

Fig. 3 Scatter plots of inconsistency indices for perturbed consistent matrices

to which extent the pairs of observations are comonotone. The Spearman index compares the way two indices rank the matrices of a fixed set of *PCM*s and reaches its maximum value 1 when two indices produce the same ranking, from the best matrix to the worst one. This means that the two indices are related by a monotone increasing function, no matter the linearity of the relation. Therefore, the Pearson Correlation Coefficient and the Spearman Rank Correlation Coefficient measure a different type of association.

Let $r(I_i(\mathbf{A}_p))$ be the rank of matrix $\mathbf{A}_p$ according to inconsistency index I_i and $r(I_j(\mathbf{A}_p))$ be the rank of matrix $\mathbf{A}_p$ according to inconsistency index I_j. For example, $r(I_i(\mathbf{A}_p)) = 1$ means that index I_i considers matrix $\mathbf{A}_p$ the most consistent one in the set. The Spearman index is defined as

$$\varrho(i, j) = 1 - \frac{6\sum_{p=1}^{N} d_p^2}{N(N^2 - 1)}, \tag{23}$$

where $d_p^2 = [r(I_i(\mathbf{A}_p)) - r(I_j(\mathbf{A}_p))]^2$.

Similarly to Sect. 4.2, we computed the Spearman Coefficient (23) for all $\binom{10}{2}$ pairs of indices $\{I_i, I_j\}$, $i = 1, \ldots, 9$; $j = i, \ldots, 10$. We report the $\varrho(i, j)$ values obtained for set S_1 in Table 3 and the values obtained for set S_2 in Table 4.

Table 1 Linear correlation computed on 10000 randomly generated *PCM*s of order 6

Index	CI	$CI^*, -c_3$	LS	GCI, ρ	HCI	I_{CD}	RE	K	GW	NI_n^σ
CI	1.	0.952	0.880	0.977	0.835	0.902	0.796	0.608	0.851	0.331
$CI^*, -c_3$	0.952	1.	0.868	0.921	0.787	0.837	0.699	0.513	0.759	0.342
LS	0.880	0.868	1.	0.870	0.707	0.843	0.590	0.474	0.733	0.496
GCI, ρ	0.977	0.921	0.870	1.	0.846	0.949	0.853	0.578	0.897	0.292
HCI	0.835	0.787	0.707	0.846	1.	0.803	0.730	0.470	0.896	0.210
I_{CD}	0.902	0.837	0.843	0.949	0.803	1.	0.778	0.447	0.833	0.276
RE	0.796	0.699	0.590	0.853	0.730	0.778	1.	0.552	0.890	0.132
K	0.608	0.513	0.474	0.578	0.470	0.447	0.552	1.	0.574	0.091
GW	0.851	0.759	0.733	0.897	0.896	0.833	0.890	0.574	1.	0.208
NI_n^σ	0.331	0.342	0.496	0.292	0.210	0.276	0.132	0.091	0.208	1

Table 2 Linear correlation computed on 10000 perturbed consistent *PCM*s of order 6

Index	CI	$CI^*, -c_3$	LS	GCI, ρ	HCI	I_{CD}	RE	K	GW	NI_n^σ
CI	1.	0.954	0.885	0.989	0.898	0.935	0.692	0.715	0.880	0.746
$CI^*, -c_3$	0.954	1.	0.807	0.909	0.838	0.839	0.602	0.571	0.756	0.625
LS	0.885	0.807	1.	0.894	0.789	0.863	0.475	0.676	0.824	0.819
GCI, ρ	0.989	0.909	0.894	1.	0.908	0.967	0.710	0.762	0.924	0.771
HCI	0.898	0.838	0.789	0.908	1.	0.888	0.670	0.675	0.924	0.656
I_{CD}	0.935	0.839	0.863	0.967	0.888	1.	0.680	0.727	0.925	0.719
RE	0.692	0.602	0.475	0.710	0.670	0.680	1.	0.625	0.710	0.480
K	0.715	0.571	0.676	0.762	0.675	0.727	0.625	1.	0.812	0.764
GW	0.880	0.756	0.824	0.924	0.924	0.925	0.710	0.812	1.	0.766
NI_n^σ	0.746	0.625	0.819	0.771	0.656	0.719	0.480	0.764	0.766	1.

4.4 Cohen's κ coefficient

The last statistical tool we used to evaluate the agreement between the inconsistency indices is the Cohen's kappa Coefficient κ (Cohen 1968). This coefficient has been designed to measure the strength of agreement in classifying processes by taking into account the agreement occurring by chance. We proceeded as follows. For each fixed inconsistency index I_i, we partitioned the set $\mathcal{A}$ of all *PCM*s in ten subsets of the same cardinality according to the inconsistency values $I_i(\mathbf{A}_p)$. As a result, index I_i classifies in the first subset the 'best' 10 % matrices, in the second subset the subsequent 10 % matrices and so on. We denote by $\{\mathcal{A}_i^1, \ldots, \mathcal{A}_i^{10}\}$ the partition of $\mathcal{A}$ induced by I_i, $i = 1, \ldots, 10$. From each partition, the thresholds values of I_i between the classes were derived. We estimated the threshold values for each inconsistency index by randomly sampling 30000 *PCM*s from $\mathcal{A}$.

We carried on the study on the agreement between the inconsistency indices based on the Cohen's kappa Coefficient similarly to the studies described in the previous subsections. First, we fixed one of the two sets of *PCM*s described above, say S_1. Then, for each inconsistency index I_i we classified the matrices in S_1 assigning them to the corresponding classes $\{\mathcal{A}_i^1, \ldots, \mathcal{A}_i^{10}\}$ according to the threshold values obtained above. Obviously, this classifying process retraces the one used for obtaining the threshold values in the partition of $\mathcal{A}$, but we

Table 3 Spearman index computed on 10000 randomly generated *PCM*s of order 6

Index	*CI*	$CI^*, -c_3$	*LS*	GCI, ρ	*HCI*	I_{CD}	*RE*	*K*	*GW*	NI_n^σ
CI	1.	0.976	0.876	0.974	0.831	0.919	0.767	0.804	0.847	0.459
$CI^*, -c_3$	0.976	1.	0.857	0.938	0.795	0.850	0.727	0.876	0.806	0.460
LS	0.876	0.857	1.	0.860	0.703	0.834	0.577	0.736	0.748	0.583
GCI, ρ	0.974	0.938	0.860	1.	0.841	0.967	0.839	0.714	0.906	0.404
HCI	0.831	0.795	0.703	0.841	1.	0.807	0.714	0.608	0.900	0.321
I_{CD}	0.919	0.850	0.834	0.967	0.807	1.	0.811	0.620	0.886	0.370
RE	0.767	0.727	0.577	0.839	0.714	0.811	1.	0.506	0.884	0.263
K	0.804	0.876	0.736	0.714	0.608	0.620	0.506	1.	0.589	0.484
GW	0.847	0.806	0.748	0.906	0.900	0.886	0.884	0.589	1.	0.338
NI_n^σ	0.459	0.460	0.583	0.404	0.321	0.370	0.263	0.484	0.338	1

Table 4 Spearman index computed on 10000 perturbed consistent *PCM*s of order 6

Index	*CI*	$CI^*, -c_3$	*LS*	GCI, ρ	*HCI*	I_{CD}	*RE*	*K*	*GW*	NI_n^σ
CI	1.	0.999	0.904	0.999	0.926	0.973	0.755	0.960	0.949	0.837
$CI^*, -c_3$	0.999	1.	0.903	0.997	0.924	0.967	0.754	0.968	0.945	0.838
LS	0.904	0.903	1.	0.903	0.790	0.878	0.505	0.866	0.855	0.928
GCI, ρ	0.999	0.997	0.903	1.	0.928	0.981	0.754	0.951	0.955	0.830
HCI	0.926	0.924	0.790	0.928	1.	0.915	0.753	0.878	0.976	0.716
I_{CD}	0.973	0.967	0.878	0.981	0.915	1.	0.735	0.899	0.955	0.787
RE	0.755	0.754	0.505	0.754	0.753	0.735	1.	0.723	0.740	0.497
K	0.960	0.968	0.866	0.951	0.878	0.899	0.723	1.	0.883	0.821
GW	0.949	0.945	0.855	0.955	0.976	0.955	0.740	0.883	1.	0.777
NI_n^σ	0.837	0.838	0.928	0.830	0.716	0.787	0.497	0.821	0.777	1.

preferred to keep the two phases distinct for procedural correctness. Finally, for all $\binom{10}{2}$ pairs of indices $\{I_i, I_j\}$, $i = 1, \dots, 9$; $j = i, \dots, 10$ we computed the Cohen's kappa Coefficient obtaining the agreement between I_i and I_j in classifying the matrices of S_1. We denote, for clarity, the obtained values by $\kappa(i, j)$. When there are more than two classes expressed on an ordinal scale (in our case we have ten classes), disagreements may not all be equally important and a greater penalty can be applied if the classes are further apart. To account for these inequalities, Cohen (1968) introduced weights in the formulation of the agreement index, leading to the *weighted* kappa coefficient

$$\kappa(i, j) = \frac{p_{o(w)} - p_{e(w)}}{1 - p_{e(w)}}. \tag{24}$$

In (24), $p_{o(w)} = \sum_{s=1}^{10} \sum_{t=1}^{10} w_{st} p_{st}$ is the observed agreement proportion, $w_{st} = 1 - \frac{(s-t)^2}{9^2}$ are the quadratic weights, and p_{st} are the observed proportions, namely, the proportion of matrices that index I_i classifies in class $\mathcal{A}_s$ and index I_j classifies in class $\mathcal{A}_t$. Finally, $p_{e(w)} = \sum_{s=1}^{10} \sum_{t=1}^{10} w_{st} p_{s\cdot} p_{\cdot t}$ is the agreement proportion expected by chance, where $p_{s\cdot}$ and $p_{\cdot t}$ are the marginal values. We reported the values $\kappa(i, j)$ obtained for the *PCM*s of S_1 in Table 5. Then, we repeated the same study on the set S_2 and we reported the results in Table 6.

Table 5 Cohen's kappa computed on 10000 randomly generated *PCM*s of order 6

Index	CI	$CI^*, -c_3$	LS	GCI, ρ	HCI	I_{CD}	RE	K	GW	NI_n^σ
CI	1.	0.966	0.865	0.965	0.820	0.909	0.757	0.795	0.836	0.458
$CI^*, -c_3$	0.966	1.	0.846	0.928	0.783	0.839	0.717	0.868	0.795	0.459
LS	0.865	0.846	1.	0.850	0.692	0.824	0.567	0.726	0.738	0.580
GCI, ρ	0.965	0.928	0.850	1.	0.828	0.958	0.829	0.706	0.894	0.405
HCI	0.820	0.783	0.692	0.828	1.	0.796	0.703	0.599	0.891	0.321
I_{CD}	0.909	0.839	0.824	0.958	0.796	1.	0.801	0.611	0.876	0.368
RE	0.757	0.717	0.567	0.829	0.703	0.801	1.	0.498	0.873	0.262
K	0.795	0.868	0.726	0.706	0.599	0.611	0.498	1.	0.579	0.484
GW	0.836	0.795	0.738	0.894	0.891	0.876	0.873	0.579	1.	0.338
NI_n^σ	0.458	0.459	0.580	0.405	0.321	0.368	0.262	0.484	0.338	1

Table 6 Cohen's kappa computed on 10000 perturbed consistent *PCM*s of order 6

Index	CI	$CI^*, -c_3$	LS	GCI, ρ	HCI	I_{CD}	RE	K	GW	NI_n^σ
CI	1.	0.905	0.580	0.805	0.540	0.513	0.067	0.569	0.603	0.037
$CI^*, -c_3$	0.905	1.	0.553	0.667	0.573	0.354	0.080	0.720	0.567	0.048
LS	0.580	0.553	1.	0.481	0.576	0.332	0.078	0.483	0.709	0.067
GCI, ρ	0.805	0.667	0.481	1.	0.445	0.756	0.041	0.313	0.593	0.020
HCI	0.540	0.573	0.576	0.445	1.	0.267	0.137	0.519	0.753	0.088
I_{CD}	0.513	0.354	0.332	0.756	0.267	1.	0.023	0.142	0.418	0.010
RE	0.067	0.080	0.078	0.041	0.137	0.023	1.	0.159	0.104	0.391
K	0.569	0.720	0.483	0.313	0.519	0.142	0.159	1.	0.419	0.104
GW	0.603	0.567	0.709	0.593	0.753	0.418	0.104	0.419	1.	0.054
NI_n^σ	0.037	0.048	0.067	0.020	0.088	0.010	0.391	0.104	0.054	1.

4.5 Single entry modification

Following Choo and Wedley (2004), we considered the consistent matrix $\mathbf{A} = (\frac{w_i}{w_j})$ corresponding to the preference vector $\mathbf{w} = (1, 2.5, 4, 5.5, 7, 8.5)$. Then, we proceeded to generate the $(n(n-1))/2 = 15$ matrices $\mathbf{C}_1, \ldots, \mathbf{C}_{15}$ by modifying a single entry of the matrix $\mathbf{A}$, $a_{ij} \to a_{ij}(1+0.9)$, $i = 1, \ldots, 5$; $j = i+1, \ldots, 6$ and putting accordingly $a_{ji} = 1/a_{ij}$. Note that this change corresponds to modifying the evaluation of a single comparison, namely the one between alternative x_i and x_j. For each one of the 15 matrices, we computed the value of the 10 inconsistency indices $I_1(\mathbf{C}_p), \ldots, I_{10}(\mathbf{C}_p)$, $p = 1, \ldots, 15$. It is interesting to note a different behavior of the indices with respect to the change of a single comparison. Five indices out of ten assign an inconsistency value which is independent from p, i.e. from the modified comparison. For example, it is $CI(\mathbf{C}_1) = CI(\mathbf{C}_2) = \cdots = CI(\mathbf{C}_{15})$ and the same holds for CI^*, GCI, I_{CD} and K. Conversely, indices LS, HCI, RE, GW and NI_n^σ assign different inconsistency values to the 15 *PCM*s. As an example, in Fig. 4 the comparison between indices CI and HCI evidences the constant value $CI(\mathbf{C}_p) \approx 0.0093$ for $p = 1, \ldots, 15$ and the different values of HCI. We performed also on the matrices $\mathbf{C}_1, \ldots, \mathbf{C}_{15}$ the study described in Sects. 4.1, 4.2, 4.3 and 4.4 for sets S_1 and S_2 but we omit, for reasons of space, to report in

Fig. 4 Scatter plot *CI–HCI*

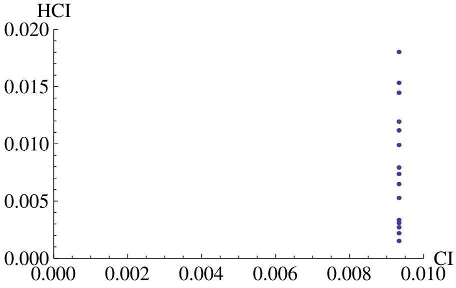

Table 7 Minimum mean value of the Spearman index

Dimension	mmv	Outlier
10	0.468	NI_n^σ
9	0.717	K
8	0.790	RE
7	0.840	

detail the corresponding results. However, we remain available to send the interested readers the tables and plots.

5 Discussion

A relevant finding emerges from the study described above. A majority of indices evidences a good internal agreement, whereas few indices perform as outliers, with poor agreement when compared with all the others. In particular, we highlight this behavior by referring to Table 3 and by operating as follows. First, we compute for each row of the table, namely for each inconsistency index, the mean value. Then, we consider the index with the minimum mean value as an 'outlier', we remove it from the table, thus obtaining a 9×9 table. We repeat the same operations on this table by identifying and removing the next outlier. We proceed the same way until for each remained inconsistency index the mean value of the Spearman index is greater than 0.8, which is considered a threshold for a good agreement. In Table 7 we report the minimum mean value of the Spearman index, say mmv, in the tables obtained as described above, for dimension 10, 9, 8 and 7. It can be checked that the indices iteratively classified as outliers are $I_{10} = NI_n^\sigma$, $I_8 = K$, and $I_7 = RE$.

Some comments are needed, for trying to interpret and justify this result. Indices NI_n^σ and K are defined using the max operator, thus focusing on a single piece of information evidencing a local maximal inconsistency. Conversely, the other indices synthesize several inconsistency contributions. Therefore, indices NI_n^σ and K could be conveniently used when it is crucial to identify the maximal violation of consistency rather than to evaluate the global amount of this violation. Nevertheless, the rather poor agreement between NI_n^σ and K evidenced in Sect. 4 remarks the diversity between the two indices.

Index $I_7 = RE$ was proposed by Barzilai (1998) and has the relevant property of being invariant w.r.t. transformation $f(a_{ij}) = a_{ij}^b$ (Fedrizzi and Brunelli 2009). This means that index RE takes into account only the mutual coherence of the judgments and it is independent from the size of a PCM's entries. Therefore, also a PCM which is very close to a consistent matrix, according to the euclidean metric, can have a large value of RE if the preferences are not coherent and, possibly, several cycles are contained. Formally, this property induces a discontinuity of *RE* in moving away from consistency. To our best knowledge, no other inconsistency index shares the same property. Moreover, Barzilai assumed that preferences can be quantified by means of every positive real number, $a_{ij} \in \mathbb{R}$. It is therefore not surprising that *RE* (16) behaves differently from the other inconsistency indices.

If we instead focus on the common characteristics of indices with a very high agreement, i.e. *GCI*, CI^*, I_{CD} and *CI*, we note that all of them, apart from *CI*, are more or less explicit averages of the single deviations of triples of pairwise comparisons from the condition of transitivity (1). Note also that, from the study performed in Sect. 4.5, all these four indices evidence the same constant behavior in changing a single comparison.

From the analysis performed on matrices of order 4 and 8 we observed that the results were coherent with those described in Sect. 4 for matrices of order 6. Nevertheless, we noted that, as the order of the matrices increases, there is a general weak decrease of the coefficients described in Sects. 4.1, 4.2, 4.3 and 4.4, namely Pearson's, Spearman's and Cohen's coefficients. In our opinion, this fact may indicate a divergent behavior of the considered inconsistency indices, thus evidencing that the problem of consistency evaluation becomes more critical as the number of involved alternatives increases. As previously specified in Sect. 4.5, we encourage the interested readers to contact us and ask for specific data which, for reasons of space, we did not include in this paper.

We conclude by stressing again that every inconsistency index is in fact a different *definition* of inconsistency degree and that there is still no recognized benchmark to measure the goodness of the various proposals.

Acknowledgements We are very grateful for the reviewers' thorough and constructive comments.

References

Aguaròn, J., & Moreno-Jimènez, J. M. (2003). The geometric consistency index: approximated threshold. *European Journal of Operational Research*, *147*(1), 137–145.

Anholcer, V., Babiy, V., Bózoki, S., & Koczkodaj, W. W. (2011). A simplified implementation of the least squares solution for pairwise comparisons matrices. *Central European Journal of Operations Research*, *19*(4), 439–444.

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

Bózoki, S. (2008). Solution of the least squares method problem of pairwise comparison matrices. *Central European Journal of Operations Research*, *16*(4), 345–358.

Brunelli, M. (2011). A note on the article "Inconsistency of pair-wise comparison matrix with fuzzy elements based on geometric mean" [Fuzzy Sets and Systems 161 (2010) 1604–1613]. *Fuzzy Sets and Systems*, *161*(1), 1604–1613. 2010, Fuzzy Sets and Systems, 176, 76–78.

Brunelli, M., Critch, A., & Fedrizzi, M. (2013). On the proportionality between some consistency indices in the AHP. *Applied Mathematics and Computation*. doi:10.1016/j.amc.2013.01.036.

Cavallo, B., & 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. (2010). Characterizations of consistent pairwise comparison matrices over abelian linearly ordered groups. *International Journal of Intelligent Systems*, *25*(10), 1035–1059.

Choo, E., & Wedley, W. (2004). A common framework for deriving preference values from pairwise comparison matrices. *Computers & Operations Research*, *31*(6), 893–908.

Chu, A. T. W., Kalaba, R. E., & Springarn, K. (1979). A comparison of two methods for determining the weights of belonging to fuzzy sets. *Journal of Optimization Theory and Applications*, *27*(4), 531–538.

Cohen, J. (1968). Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, *70*(4), 213–220.

Crawford, G., & Williams, C. (1985). A note on the analysis of subjective judgement matrices. *Journal of Mathematical Psychology*, *29*(4), 25–40.

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

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

Golden, B. L., & Wang, Q. (1989). An alternate measure of consistency. In B. L. Golden, E. A. Wasil, & P. T. Harker (Eds.), *The analytic hierarchy process, applications and studies* (pp. 68–81). Berlin: Springer.

Karelaia, N., & Hogarth, R. M. (2008). Determinants of linear judgment: a meta-analysis of lens studies. *Psychological Bulletin*, *134*(3), 404–426.

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

Linares, P. (2009). Are inconsistent decisions better? An experiment with pairwise comparisons. *European Journal of Operational Research*, *193*(2), 492–498.

Miller, G. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97.

Osei-Bryson, K.-M. (2006). An action learning approach for assessing the consistency of pairwise comparison data. *European Journal of Operational Research*, *174*(1), 234–244.

Pelàez, J. I., & Lamata, M. T. (2003). A new measure of consistency for positive reciprocal matrices. *Computers & Mathematics with Applications*, *46*(12), 1839–1845.

Phelps, R. H., & Shanteau, J. (1978). Livestock judges: how much information can an expert use? *Organizational Behavior and Human Performance*, *21*(2), 209–219.

Ramík, J., & Korviny, P. (2010). Inconsistency of pair-wise comparison matrix with fuzzy elements based on the geometric mean. *Fuzzy Sets and Systems*, *161*(1), 1604–1613.

Ramík, J., & Perzina, R. (2010). A method for solving fuzzy multicriteria decision problems with dependent criteria. *Fuzzy Optimization and Decision Making*, *9*(2), 123–141.

Saaty, T. L. (1974). Measuring the fuzziness of sets. *Journal of Cybernetics*, *4*(4), 53–61.

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

Saaty, T. L. (1980). *The analytical hierarchy process*. New York: McGraw-Hill.

Saaty, T. L. (1994). Highlights and critical points in the theory and application of the analytic hierarchy process. *European Journal of Operational Research*, *74*(3), 426–447.

Salo, A. A. (1993). Inconsistency analysis by approximately specified priorities. *Mathematical and Computer Modelling*, *17*(4–5), 123–133.

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

Shanteau, J. (1989). Psychological characteristics and strategies of expert decision makers. In B. Rohrmann, L. R. Beach, C. Vleck, & S. R. Watson (Eds.), *Advances in decision research* (pp. 203–215). Amsterdam: North-Holland.

Shanteau, J., Weiss, D. J., Thomas, R. P., & Pounds, J. (2003). How can you tell if someone is an expert? Performance-based assessment of expertise. In S. L. Schneider & J. Shanteau (Eds.), *Emerging perspectives on judgment and decision research* (pp. 620–642). Cambridge: Cambridge University Press.

Shiraishi, S., & Obata, T. (2002). On a maximization problem arising from a positive reciprocal matrix in the AHP. *Bulletin of Informatics and Cybernetics*, *34*(2), 91–96.

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

Shiraishi, S., Obata, T., Daigo, M., & Nakajima, N. (1999). Assessment for an incomplete matrix and improvement of the inconsistent comparison: computational experiments. In *Proceedings of ISAHP 1999*, Kobe, Japan.

Slovic, P. (1969). Analyzing the expert judge: a descriptive study of a stockbroker's decision processes. *Journal of Applied Psychology*, *53*(4), 255–263.

Snedecor, G. W., & Cochran, W. C. (1980). *Statistical methods*. Ames: 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., & Mizzi, P. J. (2007). The harmonic consistency index for the analytic hierarchy process. *European Journal of Operational Research*, *177*(1), 488–497.

Thurstone, L. L. (1927). A law of comparative judgement. *Psychological Review*, *34*(4), 273–386.

Weiss, D. J., & Shanteau, J. (2003). Empirical assessment of expertise. *Human Factors*, *45*(1), 104–114.

Weiss, D. J., Shanteau, J., & Harries, P. (2006). People who judge people. *Journal of Behavioral Decision Making*, *19*(5), 441–454.