PAIRWISE JUDGMENTS CONSISTENCY IMPACT ON QUALITY OF MULTI-CRITERIA GROUP DECISION-MAKING WITH AHP

Pawel Tadeusz Kazibudzki, Jiří Křupka

Introduction

Presumably complex systems can be better understood when they are broken down into their constituent elements and structured hierarchically. Then, judgments about these elements can be synthesized on the basis of their relative importance at each level of the hierarchy into a set of overall priorities. By breaking down a reality into homogenous clusters and subdividing them into smaller ones, it is possible to integrate large amounts of information into the structure of a problem and form a more comprehensive picture of the whole system.

There is a decision support methodology (DSM) which conforms to the above prescription. It is called the Analytic Hierarchy Process (AHP) and was devised at the Wharton School of Business by Thomas Saaty (1980). Its contemporary applications can be found, for example in Lidinska and Jablonsky (2018), Abdelmaguid and Elrashidy (2016), Kramulová and Jablonský (2016), and Ponis et al. (2015). This DSM is based on the pairwise judgments technique which comes from an influential paper of Marquis de Condorcet (1785), who used this technique in the election process (Young, 1988), and which was popularized by Thurstone (1927) thanks to his first contemporary application at the beginning of the 20th century.

When group decision-making (GDM) is taken into consideration, the AHP seems a particularly attractive methodology, and although it has been examined numerous times from the perspective of its effectiveness and applicability in GDM processes (see e.g. Scala et al., 2016; Saaty & Peniwati, 2008; Saaty & Vargas, 2012; Aguarón et al., 2014; Hosseinian et al., 2012; Moreno-Jiménez et al., 2005, 2008; Altuzarra et al., 2010; Sun & Greenberg, 2006), still a research gap could have been identified. Thus, this paper examines judgments

consistency influence on the credibility of priority ratios (PRs) within a particular priority vector (PV) derived from inconsistent pairwise judgments made by a decision maker (DM). Examination results generalize to the synthesized pairwise comparison matrix that is obtained on the basis of individual pairwise comparison matrices for all group members. Having in mind that a consistency index for the PCM denoting group preferences cannot be greater than the consistency index of the most inconsistent individual PCM it became possible to designate the credibility of the priority vector for the group on the basis of the most inconsistent individual PCM.

The article is organized around three main sections: the introductory section which elaborates on pairwise judgments, AHP, and GDM with application of AHP; the methodological section devoted to the research methodology, comprising an illustrative example of the problem and selected pitfalls during priority ratios estimation process which builds on preselected measures of estimation errors; the investigational section encompassing the research outcome, its contribution to the research field and the examination breakthrough from the viewpoint of other research papers. The final part of the article constitutes the section 'Conclusions' which summarizes examination findings.

1. Background

The AHP can be considered to be both a descriptive and prescriptive model of decision making. It promotes pairwise judgments (i.e. valuation on the basis of pairwise comparisons) of criteria and alternatives with respect to a criterion. Genuinely (as proposed by the creator of AHP), the comparison process proceeds with the application of a fundamental scale of absolute numbers that has been

proven in practice i.e. Saaty's numerical scale which comprises of the integers from one (equivalent to the verbal judgment - 'equally preferred') to nine (equivalent to the verbal judgment - 'extremely preferred'), and their reciprocals. Other numerical scales have been considered also see e.g. Dong et al. (2008). The methodology of AHP is based on the well-defined mathematical structure of consistent matrices and their associated principal right eigenvector's (REV) ability to generate true or approximate weights, see e.g. Merkin (1979), Saaty and Vargas (1984).

Generally, the problem of deriving PRs from a pairwise comparison matrix (PCM) denoted as $A = [a_{ij}]_{nxn}$ with elements $a_{ij} = a_i/a_j$, is to estimate $w = [w_1, w_2, w_3, …, w_n]^T$ on the basis of matrix A which comprises a decision maker's pairwise judgments (denoting DM preferences) concerning the importance of a given binary set of alternatives. Commonly PRs w_i, where $I = 1,…, n$, are selected to be positive and normalized to unity $\sum_i^n w_i = 1$, and the elements a_{ij} of matrix A are then the DM's judgments about the PRs $w_{ij} = w_i/w_j$, where $i, j = 1,…,n$, and n is the number of all alternatives being considered. In a perfect judgment case then, the problem can be designated as:

$$A \cdot w = \lambda \cdot w \qquad (1)$$

and w can be computed by solving the eigenvector equation (1). In a perfect case (matrix A is consistent) λ is the only nonzero eigenvalue of A i.e. the nonzero solution of the characteristic equation:

$$\det(A - \lambda \cdot I) = 0 \qquad (2)$$

where I denotes the identity matrix of order n. In this case, also $\lambda = n$. On the other hand, when the case is not perfect (matrix A is not consistent) an estimate of the true w is the normalized principal right eigenvector (REV) associated with the maximal eigenvalue. Thus, in order to obtain the estimate it is needed to solve the general eigenvector equation:

$$A \cdot w = \lambda_{max} \cdot w \qquad (3)$$

where λ_{max} denotes the principal eigenvalue which is not smaller than n, is simple and its existence is guaranteed by the Perron-Frobenius Theorem, see e.g. Saaty and Vargas (1984).

The matrix of ratios $A = (w_i/w_j)$ is consistent, if and only if n is its principal eigenvalue and $A \cdot w = n \cdot w$. Further, $w > 0$ is unique to within a multiplicative constant.

If the elements of a matrix A satisfy the condition $w_{ij} = 1/w_{ji}$ for all $i, j = 1,…, n$ then the matrix A is said to be *reciprocal*. If its elements satisfy the condition $w_{ik}w_{kj} = w_{ij}$ for all $i, j, k = 1,…, n$ and the matrix is *reciprocal*, then it is called *cardinally transitive* or *consistent*. Matrix A can also be only *transitive* if the following conditions hold: (i) if for any $i = 1,…, n$, an element w_{ij} is not less than an element w_{ik} then $w_{ij} \geq w_{ik}$ for $i = 1,…, n$, and (ii) if for any $i = 1,…, n$, an element w_{ji} is not less than an element w_{ki} then $w_{ji} \geq w_{ki}$ for $i = 1,…, n$. In the case of reciprocal PCMs – which are the only accepted PCMs for the AHP although counterarguments exist in literature (see e.g. Linares et al., 2016) the two conditions (i) and (ii) are equivalent.

Fundamentally, all theories are based on axioms, so is the AHP. Its creator Saaty (2006) defines five conditions for good approximations: reciprocity, homogeneity (the elements being compared must be of the same order of magnitude), independency (judgments about, or the priorities of, the elements in a hierarchy cannot depend on lower level elements), near consistency and uniform continuity (elements w_i, $I = 1,…, n$ should be relatively insensitive to small changes in the elements a_{ij}, only then good approximations to a_{ij} remain w_i/w_j ratios).

The central point of AHP and the key issue for a theory of choice that is based upon AHP is the methodology of capturing (in)consistency of PCMs within the AHP. In order to derive credible priority vectors (PV) within AHP, it is necessary to impose some boundaries on (in) consistency of PCMs involved in the process. Indeed, significant violation of the PCM (in) consistency may mislead the true values of priority ratios within the PV making the entire methodology itself useless. On the other hand, it does not mean that a high consistency of PCM guarantees credible values of PV because even perfectly consistent PCMs may not be error free, see e.g. Grzybowski (2016) and Temesi (2011). That is why establishing some relations between (in)consistency of PCM and the credibility of priority ratios estimates seems so important.

The AHP genuine measure of PCM (in) consistency belongs to Saaty (1980) and is strictly related to the REV, which makes it

especially attractive. It does not mean that the other PCM inconsistency measures (called consistency or inconsistency indices) do not exist. To the contrary, a number of other indices can be found in literature, see e.g. Mizuno (2019), Peláez, Martínez and Vargas (2018), Dixit (2018), Fedrizzi and Ferrari (2017), proposed quite recently.

A detailed analysis of all consistency indices available in literature is beyond the scope of this research. However, a reader interested in various approaches to consistency measurement during pairwise comparisons may want to review those references. It behooves to mention that pairwise judgments consistency measurement was also a topic of more cross-sectional surveys e.g. Brunelli (2018), Kou et al. (2016). Having the perspective on a scale of research devoted to various ways of pairwise comparisons consistency identification, it remains to mention that Saaty's concept for pairwise comparisons consistency measurement proposed for AHP is currently systematically criticized, see e.g. Xu et al. (2008), Koczkodaj and Szwarc (2014), Koczkodaj and Urban (2018).

However, taking into account that the AHP creator's concept is still applied in the way it was proposed a few decades ago, interested readers in a detailed perspective of Saaty's concept, as well a more fundamental analysis of the whole AHP approach, may want to study a more detailed examination of this methodology for instance in Wu and Kou (2016), Kou et al. (2016), and Saaty (2008b). Further discussion within this area, for reasons of brevity, is deliberately omitted. Instead, some key issues of group decision making with the application of AHP will be briefly depicted.

It is a fact that most of real-life decisions are not made by individuals, but by groups of individuals e.g. committees, councils, etc. From that perspective the relation between the quality of pairwise judgments made by individuals and the quality of the representative judgment for a group of individuals is of great importance.

Thus, the prescription exists for an individual judgments aggregation in a way which enables obtaining a representative group judgment. The reciprocal property of the AHP plays an important role from that perspective. Generally, judgments have to be combined in such a way that reciprocals of the synthesized judgments are equal to the syntheses of these judgment reciprocals.

It has been deduced that the only unique way to do that is to apply the geometric mean procedure. It can be done in two ways (Saaty, 2008b):

- if experts are appointed as decision makers, then rather than combining their individual judgments, their final outcome from a hierarchy is synthesized with application of a geometric mean;
- on the other hand, if the individuals themselves have different degrees of importance i.e. voting powers, their individual judgments are raised to their voting power and the group outcome is established on the basis of their individual judgments i.e. the weighted geometric mean is formed (Formulae 7 and 8).

$$\overline{a_{ij}} = \left(\prod_{k=1}^{n} a_{ijk}^{w_k} \right)^{\left(\sum_{k=1}^{n} w_k \right)^{-1}} \tag{7}$$

$$\overline{a_{ij}} = \exp\left(h \prod_{k=1}^{n} a_{ijk}^{w_k} \middle/ \sum_{k=1}^{n} w_k \right) \tag{8}$$

where w_k denotes a priority of importance for the individual.

In the latter case, the final outcome for a hierarchy is computed with the application of the standard AHP aggregation i.e. with application of the weighted arithmetic mean (the priority of the particular alternative under its criterion is weighted by the priority of its criterion(s), then the total priority of the given alternative is determined by the sum of their weighted priorities).

It has been proven that application of the geometric mean procedure for individual preferences aggregation is the only one which satisfies a number of important properties (Aczel & Saaty, 1983). It has been also proven, (see e.g. Liu, Zhang, & Wang, 2012; Grošelj & Stirn, 2012; Escobar, Aguarón, & Moreno-Jimenez, 2004; Xu, 2000), that the $CI(A)$ of the group preferences cannot be greater than the $CI(A^*)$ of the most inconsistent individual PCM = A^*, i.e.:

$$CI(A) \leq \max\{CI(A_1), CI(A_2),...CI(A_n)\}.$$

However, despite the relevance of the findings stated above, it needs to be

stressed that relatively consistent individual pairwise judgments do not guarantee 100% preferences credibility derived thereof, see e.g. Temesi (2011). Thus, rather then focusing on inconsistency of the group preferences, primarily the consistency of individual pairwise judgments must be meticulously controlled from the perspective of its relationship with priority ratios estimation errors which can distort credibility of a particular priority vector. The necessity of research in this area seems to be paramount.

2. Research Methodology

It is emphasized, that few research papers have dealt in depth with the above presented problem i.e. the relation between a level of the pairwise judgments (in)consistency and the range of possible estimation errors for established priority ratios, see e.g. Grzybowski (2016) and Kazibudzki (2019a).

However, despite of the relevance of findings published therein, those research studies concentrate on average estimation errors within particular priority vectors. The consequence of such a perspective is a tacit assumption that estimation errors for particular priority ratios are more or less the same as the mean error for the particular priority vector. It turns out, that it is not necessarily true, especially when relative measures for estimation errors are considered. The examination of these issues is in order and will be made briefly in the below subsection entitled 'Problem illustration', and then thoroughly in the paper's subsection entitled 'Examination breakthrough'. The problem exemplification is first to be considered.

2.1 Problem Illustration

The following hypothetical normalized priority vector (the vector of priority ratios) is considered: $PV(w)$ = [0.0625, 0.1042, 0.1458, 0.1875, 0.2292, 0.2708]. It is assumed that the vector reflects 'true' (not estimated) the DM's relative preference toward six objects whose relative characteristics are known e.g. the strength of preferences toward the objects is associated with the size of these objects (their mass, volume, circumference etc.), so their relative importance can be calculated by dividing the particular object's size by the total size of all objects. On the basis of this 'true' $PV(w)$, the $PCM(w)$ is formed denoted as $A(w)$ with elements $w_{ij} = w_i / w_j$:

$$A(w) = \begin{bmatrix} 1 & 0.6 & 0.4286 & 0.3333 & 0.2727 & 0.2308 \\ 1.6667 & 1 & 0.7143 & 0.5556 & 0.4545 & 0.3846 \\ 2.3333 & 1.4 & 1 & 0.7778 & 0.6364 & 0.5385 \\ 3 & 1.8 & 1.2857 & 1 & 0.8182 & 0.6923 \\ 3.6667 & 2.2 & 1.5714 & 1.2222 & 1 & 0.8462 \\ 4.3333 & 2.6 & 1.8571 & 1.4444 & 1.1818 & 1 \end{bmatrix}$$

Then, $A(w)$ is perturbed by perturbation factor e which single value in this example is given e = 0.5. This technique allows to emulate inconsistency during DMs judgments concerning objects and is widely accepted for this purpose since its first application i.e. Zahedi (1986). In this way the perturbed matrix $A(x)$ is obtained, where $x_{ij} = w_{ij}e$ for $i \neq j$ and $i, j \in N = \{1,…,6\}$.

$$A(x) = \begin{bmatrix} 1 & 0.3 & 0.2143 & 0.1667 & 0.1364 & 0.1154 \\ 0.8333 & 1 & 0.3571 & 0.2778 & 0.2273 & 0.1923 \\ 1.1667 & 0.7 & 1 & 0.3889 & 0.3182 & 0.2692 \\ 1.5 & 0.9 & 0.6429 & 1 & 0.4091 & 0.3462 \\ 1.8333 & 1.1 & 0.7857 & 0.6111 & 1 & 0.4231 \\ 2.1667 & 1.3 & 0.9286 & 0.7222 & 0.5909 & 1 \end{bmatrix}$$

Further, the upper triangle elements of $A(x)$ are rounded to the closest value of Saaty's scale and reciprocity is imposed (only the upper triangle elements i.e. elements above $A(x)$ diagonal are considered for rounding while the lower triangle elements are computed as reciprocals of the upper triangle elements). In this way the scaled and reciprocal $A(v)$ is obtained which reflects DMs judgments concerning objects expressed with application of the particular preference scale, in this example Saaty's scale. Other scales can be applied also, for references see e.g. Dong et al. (2008).

$$A(v) = \begin{bmatrix} 1 & 0.3333 & 0.2 & 0.1667 & 0.1429 & 0.1111 \\ 3 & 1 & 0.3333 & 0.25 & 0.25 & 0.2 \\ 5 & 3 & 1 & 0.3333 & 0.3333 & 0.25 \\ 6 & 4 & 3 & 1 & 0.3333 & 0.3333 \\ 7 & 4 & 3 & 3 & 1 & 0.5 \\ 9 & 5 & 4 & 3 & 2 & 1 \end{bmatrix}$$

It should be emphasized that $A(v)$, on the basis of Saaty's consistency philosophy, should be considered as acceptably consistent because its $CI(A(v))$ = 0.0670, $RI(6)$ = 1.24 thus $CR(A(v))$ = 0.0541 which informs of

an acceptable level of consistency (Saaty suggested that CR < 0.1).

On the basis of $A(v)$, PV(v) is calculated, in this example, with the application of the REV method as the genuine AHP prioritization technique (PT). It behooves to mention that other PTs, which were suggested in literature for this purpose, can be also applied. For brevity, they will not be discussed in this research. However, the interested reader may want to find references where these PTs are scrutinized. The most recent are Orbán-Mihálykó et al. (2017), Kazibudzki (2016b), Kułakowski (2015).

Continuing the main stream of the research, on the basis of $A(v)$ the following PV(v) is obtained with application of the REV method: PV(v) = [0.0277, 0.0566, 0.1027, 0.1714, 0.2679, 0.3736], which is different than PV(w) = [0.0625, 0.1042, 0.1458, 0.1875, 0.2292, 0.2708]. Having those two vectors of priority ratios, it is possible to compute deviations among their elements i.e. maximal absolute deviation (*MaxAD*), mean absolute deviation (*MAD*), minimal absolute deviation (*MinAD*), maximal relative deviation (*MaxRD*), mean relative deviation (*MRD*), and minimal relative deviation (*MinRD*), see formulae in Tab. 1.

From the perspective of this research, devoted to designating the priority ratios estimate credibility, four deviations presented in Tab. 1 i.e. *MaxAD*, *MaxRD*, *MAD* and *MRD*, become especially significant because they enable designation of confidence intervals for 'true' priority ratios (as in the classic statistical estimation theory).

In this exclusively illustrative example, the confidence for those intervals is purely hypothetic because it cannot be designated only on the basis of one case. However, iterations of similar cases are possible using Monte Carlo simulations which results can provide meaningful data in this matter. This issue is scrutinized further in the article's subsection 'Examination methodology'.

Returning to the problem's illustration issue, the values of four deviations especially significant for the considered problem study are presented in Tab. 2. On the basis of these values (Tab. 2), the hypothetic confidence intervals for priority ratios estimates (PRE) can be established and illustrated (Fig. 1–2). As can been noticed (Fig. 1–2), the hypothetic confidence intervals for priority ratios have different features i.e. range and symmetry in relation to the estimated priority ratios values. Those features depend upon applied deviation.

Tab. 1:   **Formulae for deviations among 'true' and estimated vectors of priority ratios**

$$MaxAD(w,v) = \max\left\{\left|w_i - v_i\right|\right\}_{i=1...,n} \qquad MaxRD(w,v) = \max\left\{\frac{\left|w_i - v_i\right|}{w_i}\right\}_{i=1...,n}$$

$$MAD(w,v) = \frac{1}{n}\sum_{i=1}^{n}\left|w_i - v_i\right| \qquad MRD(w,v) = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|w_i - v_i\right|}{w_i}$$

$$MinAD(w,v) = \min\left\{\left|w_i - v_i\right|\right\}_{i=1...,n} \qquad MinRD(w,v) = \min\left\{\frac{\left|w_i - v_i\right|}{w_i}\right\}_{i=1...,n}$$

Source: own

Tab. 2:   **Deviations among 'true' and estimated vectors of priority ratios**

MAD	MaxAD	MRD	MaxRD
0.0472	0.1028	0.3239	0.5568

Source: own

Noticeably, a higher spread of hypothetic confidence intervals and their higher asymmetry is observed when relative and/or maximum deviations are applied. For relative deviations, the spread of hypothetic confidence intervals also depends on the value of the particular priority ratio i.e. higher values of priority ratios entail a larger spread for their hypothetic confidence intervals: the problem is clearly visible for *MaxRD* (Fig. 2). It is very important to notice that although the spread and asymmetry of the hypothetic confidence intervals for relative deviations are significant, the hypothetic confidence intervals for the first and second priority ratio i.e.

$$v_1 \in [0.0209, \ 0.041] \ , \ v_2 \in [0.0427, \ 0.0837],$$

established with the application of *MRD* do not encompass the 'true' values of the first and second priority ratio which equal respectively $x_1 = 0.0625$, $x_2 = 0.1042$.

It bears mentioning that the *MRD* has also another very unattractive feature i.e. it can mask significant dispersion among particular priority ratios deviation.

For example, let two normalized hypothetic vectors of priority ratios be given as PV(z) = [0.2262, 0.2729, 0.1390, 0.3619], and its estimate PV(s) = [0.2143, 0.3571, 0.1429, 0.2857]. In this case PV(z) is the 'true' priority vector in relation to which a mean relative

Hypothetic confidence intervals for PRE set with application of *MAD* and *MaxAD*
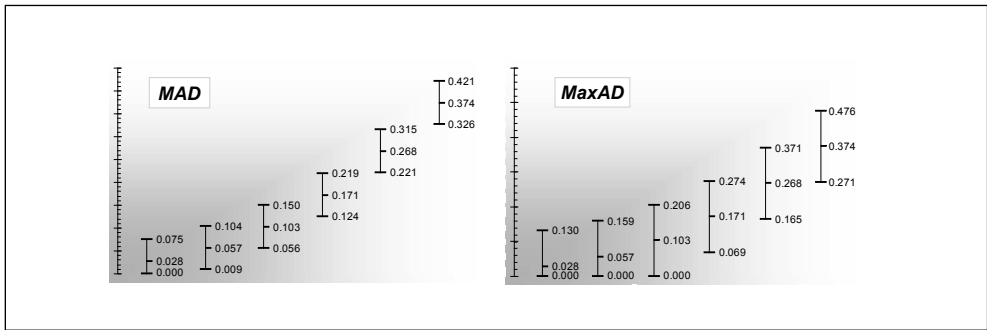
Source: own

Hypothetic confidence intervals for PRE set with application of *MRD* and *MaxRD*
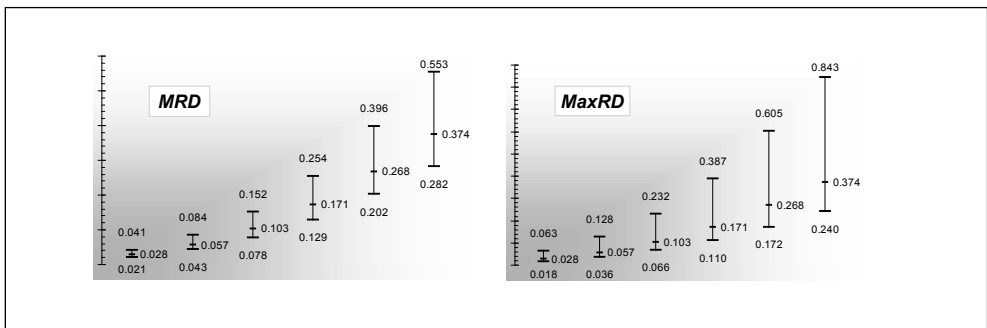
Source: own

deviation for priority ratios is calculated. In such a situation, one receives *MRD* = 0.15 which seems a rather small relative deviation. However, as can be noticed, singular relative deviations (*SRD*) among priority ratios within PV(z) and PV(s) equal, respectively *SRD* = [0.0527, 0.3087, 0.0280, 0.2106], and are highly divergent. In consequence, a reversal of priority ratios ranking is noted i.e. PV(z) = {3, 2, 4, 1} and PV(s) = {3, 1, 4, 2}. This is the exemplary situation which needs prevention, thus it is argued to withdraw relative deviations from further application to similar problems.

Nevertheless, because the above proposition is based only on a hypothetic illustrative example, it is thoroughly examined further in this paper in the subsection entitled 'Examination breakthrough' for more credible conclusions.

2.2 Examination Methodology

Continuing the main stream of the research, in real AHP applications, the 'true' vector of priority ratios (in the illustrative example denoted as PV(w)) is unknown. The entire AHP concept assumes it can be estimated with the application of the selected prioritization technique (PT), which in classic AHP is the REV method described earlier in this paper.

As can be noticed from the earlier provided example, an estimate of the unknown PV can be more or less credible. In general, this credibility generally depends on the applied preference scale, PT, and the consistency of PCM on the basis of which the estimate of unknown PV is derived. An examination concerning differences between various preference scales and PTs in relation to credibility of PVs obtained with their application is beyond the scope of this research. However, the relation between the consistency of PCM and the PV estimate credibility seems particularly attractive. The examination proceeds with the application of Monte Carlo simulations coded and performed in Wolfram Mathematica Software.

Taking into account the fact that Saaty's concept of PCM consistency measurement was seriously questioned (see e.g. Xu et al., 2008; Grzybowski, 2012 Koczkodaj & Szwarc, 2014; Grzybowski, 2016; Koczkodaj & Urban, 2018), and the credibility of REV as the PT is slightly undermined (see e.g. Kazibudzki, 2019b; Bana e Costa & Vansnick, 2008; Schoner & Wedley,

1989; Budescu et al., 1986; Belton & Gear, 1983; Johnson et al., 1979), for the Monte Carlo simulations in this research, the Logarithmic Least Squares Method (LLSM) developed by Crawford and Williams (1980, 1985) is applied, as the oldest alternative for the REV (Formulae 9 and 10), as well LLSM based consistency index $CI_{(LLSM)}$ proposed by the same authors (Formula 11), and examined by Aguarón and Moreno-Jimenez (2003).

$$w_{(LLSM)} = \min \sum_{i=1}^{n} \sum_{j=1}^{n} h^2\left(a_j \frac{w_j}{w_i}\right) \qquad (9)$$

$$w_{i(LLSM)} = \left(\prod_{j=1}^{n} a_j\right)^{1/n} \Bigg/ \sum_{i=1}^{n}\left(\prod_{j=1}^{n} a_j\right)^{1/n} \qquad (10)$$

$$CI_{(LLSM)} = \frac{2}{(n-1)(n-2)}\sum_{i<j}\log^2\left(\frac{a_j w_j}{w_i}\right) \qquad (11)$$

To properly examine the problem from the given perspective, the following simulation scenario is considered. It behooves to mention that its assumptions come from Grzybowski (2016) who first devised its framework for a similar analysis. Thus, the following steps in the scenario are considered:

Step 1: For the assumed *n*, randomly generate a 'true' [*n* × 1] priority vector $w = [w_1,\ldots, w_n]^T$ and the corresponding 'genuine' PCM(*w*) = *G*(*w*).

Step 2: Randomly select an element w_{xy} for *x* < *y* of *G*(*w*), and replace it with $w_{xy}e_B$, where e_B is a relatively significant error, randomly drawn (with application of *uniform* distribution) from the interval $e_B \in [2;4]$. Errors of that magnitude are basically considered as relatively "significant", see e.g. Dijkstra (2013), Grzybowski (2016).

Step 3: For every element w_{ij}, *i* < *j* ≤ *n*, other than w_{xy}, randomly select a value e_{ij} for the relatively small error in accordance with the given probability distribution π (applied in equal proportions as *gamma*, *log-normal*, *truncated normal*, and *uniform* distribution) and replace the element w_{ij} with the element $w_{ij}e_{ij}$ where e_{ij} is randomly drawn from the interval $e_{ij} \in [0,5;1,5]$ with application of *uniform* distribution.

Step 4: For all *i*, *j* such that *i* < *j*, round all values of $w_{ij}e_{ij}$ of *G*(*w*) to the closest value from the selected scale.

Step 5: Replace all elements w_{ij} for $i > j$ of $G(w)$ with $1/w_{ij}$. The perturbed PCM(w) in Steps 2–5 denote as $P(v)$.

Step 6: On the basis of $P(v)$ compute the value of the examined consistency index CI as well as the estimate of the vector w denoted as v derived from $P(v)$ with application of assigned prioritization technique. Then calculate $MaxAD_1$ and $MaxAD_2$ i.e. the maximum and the second maximum absolute deviation between priority ratios of w and v in accordance with formula presented in Tab. 1. Save the values computed in this step as one record.

Step 7: Repeat Steps 2–6 N_P times.

Step 8: Repeat the scenario N_T times.

Step 9: Save *all* the records as one database file.

The above presented simulation framework enables examination of relations between performance of a given consistency index and greatest deviations between a 'true' and estimated vector of priority ratios. This way it is possible to associate various values of a given consistency index with highest potential estimation errors for obtained priority ratios.

For formality, the simulation framework presented above exactly emulates steps scrutinized in the example provided earlier in this paper in subsection 'Problem illustration'. All parameters of the applied probability distributions in the simulation framework i.e. *gamma*, *log-normal*, *truncated normal*, and *uniform*, are set in such a way that the expected value $EV(e_{ij}) = 1$. In this way the simulation examination and its results reflect the reasonable assumption concerning human nature i.e. decision makers judgments are more or less deviated from optimal outcome but 'close' to it.

3. Results and Discussion

For brevity it was decided to scrutinize the examination results for $n = 4$. It behooves to mention that for $n = 3$, direct interrelation between consistency indices is observed, see e.g. Bozóki and Rapcsák (2008) and/or Dijkstra (2013).

3.1 Research Outcome

The simulation results are presented in Tab. 3 and 4. They are based on $N_P = 100$, and $N_T = 1000$.

3.2 Examination Contribution

Having the empirical distribution of maximal absolute deviations between 'true' and estimated priority vectors, the empirical confidence intervals for particular priority ratios can be established e.g. with the application of 'average maximum absolute deviation' established during simulations. On the basis of selected statistics, the credibility of the priority vector can also be designated with the selected rank of a quantile. Thus, one can expect a confidence interval with an average level of certainty for maximal absolute deviation when the average maximum absolute deviation is applied. In addition, one can expect a confidence interval noted by the rank of the quantile, when quantiles of maximum absolute deviations are applied. Noticeably, confidence intervals established on the basis of quantiles will be slightly exaggerated because the maximum absolute deviation among given priority ratios within two priority vectors cannot repeat itself (remaining deviations must be smaller). That is why another approach is proposed.

Noticeable, in multicriteria decision making processes, a decision maker (DM) is usually interested in the most attractive alternative. So, the probability of the highly ranked alternative reversal is of great importance. Thus, it is proposed to apply a maximum absolute deviation and the second maximum absolute deviation for examination, if the risk of rank reversal for the first two alternatives exists, and for examination purposes, how high the risk is.

To exemplify, the following hypothetic normalized vector of priority ratios is considered: $PV(v) = [0.62, 0.24, 0.1, 0.04]$. The $PV(v)$ designates the following ranks for evaluated options: $A_1 \prec A_2 \prec A_3 \prec A_4$. It is assumed that the $PV(v)$ was derived from the PCM for which $CI_{(LLSM)} = 0.109736$. In this case, a decision maker may wonder about the probability of the highly ranked option reversal. In light of the research outcome, the answer for this inquiry depends on the level of certainty assumed by a decision maker.

When this level equals 95%, then the 0.95-quantile of the maximum absolute deviation distribution for $CI_{(LLSM)} = 0.109736$ equals 0.208885 (Tab. 3), and 0.95-quantile of the second maximal absolute deviation distribution equals 0.150579. Thus, if a difference between the first two priority ratios of the hypothetic $PV(v)$ is higher than 0.150579 + 0.208885 i.e. 0.359464

Tab. 3:	Distribution of *MaxAD₁* i.e. the maximal absolute deviations for estimated priority ratios in relation to performance of the consistency index CI_{LLSM}						

i	*i*th interval for CI_{LLSM}	Average CI_{LLSM} within *i*th interval	*p*–quantiles of *MaxAD₁* between *w* and *v* for *i*th interval of CI_{LLSM}					Average *MaxAD₁* between *w* and *v*
			p = 0.8	*p* = 0.9	*p* = 0.95	*p* = 0.98	*p* = 0.99	
1	[0.0, 0.0202)	0.009789	0.055462	0.081919	0.123431	0.194786	0.241852	0.042064
2	[0.0202, 0.081)	0.050333	0.083198	0.127897	0.173985	0.230384	0.266555	0.058667
3	[0.081, 0.141)	0.109736	0.128978	0.167430	0.208885	0.257923	0.292974	0.086182
4	[0.141, 0.201)	0.170316	0.135142	0.175823	0.215673	0.260070	0.291170	0.097059
5	[0.201, 0.261)	0.230957	0.139692	0.181055	0.217441	0.260860	0.294106	0.100639
6	[0.261, 0.322)	0.290961	0.148930	0.185320	0.216985	0.258141	0.293103	0.104673
7	[0.322, 0.382)	0.351898	0.152956	0.183949	0.214520	0.257819	0.291483	0.108191
8	[0.382, 0.442)	0.411428	0.154353	0.184168	0.216551	0.261614	0.295555	0.111088
9	[0.442, 0.503)	0.472071	0.151550	0.181402	0.215974	0.259817	0.291782	0.111199
10	[0.503, 0.563)	0.532404	0.150774	0.184129	0.220709	0.265578	0.302564	0.111745
11	[0.563, 0.623)	0.591957	0.151667	0.187019	0.224346	0.271002	0.304387	0.112368
12	[0.623, 0.684)	0.652322	0.152885	0.189522	0.229630	0.275033	0.309352	0.112625
13	[0.684, 0.744)	0.713164	0.156414	0.196321	0.236259	0.286920	0.318705	0.114256
14	[0.744, 0.804)	0.773112	0.159058	0.201239	0.240192	0.290268	0.326163	0.115300
15	[0.804, 0.865)	0.833323	0.161285	0.202429	0.240903	0.290718	0.326226	0.116094
16	[0.865, 0.925)	0.893977	0.160628	0.203376	0.242114	0.294927	0.334319	0.116101
17	[0.925, 0.985)	0.954481	0.167689	0.210992	0.248777	0.297410	0.334572	0.119510
18	[0.985, 1.046)	1.014420	0.171211	0.215302	0.257097	0.307838	0.342969	0.120865
19	[1.046, 1.106)	1.075330	0.175970	0.219679	0.259016	0.312604	0.352029	0.122409
20	[1.106, 1.166)	1.135290	0.177132	0.222577	0.262351	0.312678	0.345093	0.123456
21	[1.166, 1.226)	1.194820	0.183751	0.229224	0.270164	0.324909	0.374050	0.128275
22	[1.226, 1.287)	1.256030	0.179100	0.228531	0.269188	0.317348	0.356733	0.125110
23	[1.287, 1.347)	1.316700	0.184083	0.229390	0.271408	0.324651	0.369775	0.128810
24	[1.347, 1.407)	1.376290	0.192343	0.237683	0.275558	0.337192	0.369280	0.131952
25	[1.407, 1.468)	1.437290	0.193035	0.244625	0.285832	0.336465	0.377465	0.133709
26	[1.468, 1.528)	1.497320	0.199676	0.241012	0.285629	0.345064	0.388863	0.136678
27	[1.528, 1.588)	1.557600	0.197001	0.245495	0.291895	0.348678	0.381236	0.134634
28	[1.588, 1.649)	1.618450	0.202120	0.252202	0.299193	0.359296	0.400384	0.138595
29	[1.649, 1.709)	1.678350	0.198996	0.242366	0.282109	0.342629	0.387303	0.140573
30	[1.709, oo)	2.494600	0.241975	0.296090	0.340766	0.392593	0.431660	0.164883

Source: own

Note: The results were generated for *n* = 4 on the basis of the presented simulation framework. The outcome is based on 100,000 perturbed reciprocal PCMs. The simulation scenario assumed LLSM as the PT and Saaty's preference scale.

Tab. 4: Distribution of *MaxAD₂* i.e. the second maximal absolute deviations for estimated priority ratios in relation to performance of the consistency index CI_{LLSM}

i	i^{th} interval for CI_{LLSM}	Average CI_{LLSM} within i^{th} interval	p–quantiles of *MaxAD₂* between w and v for i^{th} interval of CI_{LLSM}					Average *MaxAD₂* between w and v
			$p = 0.8$	$p = 0.9$	$p = 0.95$	$p = 0.98$	$p = 0.99$	
1	[0.0, 0.0202)	0.009789	0.038698	0.055876	0.083513	0.133809	0.169246	0.029146
2	[0.0202, 0.081)	0.050333	0.057643	0.088766	0.122063	0.164735	0.193358	0.041093
3	[0.081, 0.141)	0.109736	0.089991	0.119558	0.150579	0.190940	0.218659	0.060380
4	[0.141, 0.201)	0.170316	0.095839	0.124312	0.156293	0.196901	0.224054	0.067589
5	[0.201, 0.261)	0.230957	0.097189	0.128856	0.160173	0.198844	0.223367	0.070123
6	[0.261, 0.322)	0.290961	0.102280	0.134651	0.163703	0.197659	0.221966	0.073105
7	[0.322, 0.382)	0.351898	0.107486	0.136816	0.163401	0.195838	0.224857	0.075819
8	[0.382, 0.442)	0.411428	0.109694	0.137829	0.163587	0.199636	0.226573	0.077957
9	[0.442, 0.503)	0.472071	0.107667	0.134538	0.158997	0.196299	0.224823	0.077769
10	[0.503, 0.563)	0.532404	0.107582	0.135338	0.163117	0.205157	0.234390	0.078694
11	[0.563, 0.623)	0.591957	0.107866	0.135928	0.166899	0.206786	0.235808	0.079149
12	[0.623, 0.684)	0.652322	0.108464	0.138907	0.172721	0.214684	0.245777	0.079735
13	[0.684, 0.744)	0.713164	0.111545	0.143986	0.176468	0.223528	0.254954	0.081337
14	[0.744, 0.804)	0.773112	0.114840	0.145797	0.179305	0.225435	0.256358	0.082506
15	[0.804, 0.865)	0.833323	0.115525	0.149092	0.183596	0.226972	0.259410	0.083318
16	[0.865, 0.925)	0.893977	0.116169	0.149262	0.185559	0.229245	0.259710	0.083580
17	[0.925, 0.985)	0.954481	0.120943	0.154683	0.188189	0.229371	0.261401	0.085939
18	[0.985, 1.046)	1.014420	0.123626	0.158677	0.196483	0.240583	0.274190	0.087317
19	[1.046, 1.106)	1.075330	0.127007	0.161803	0.199002	0.242274	0.279624	0.088264
20	[1.106, 1.166)	1.135290	0.128787	0.164871	0.201893	0.241767	0.271120	0.089477
21	[1.166, 1.226)	1.194820	0.131613	0.169470	0.207377	0.248484	0.280735	0.092908
22	[1.226, 1.287)	1.256030	0.130773	0.169210	0.203081	0.254839	0.290318	0.090748
23	[1.287, 1.347)	1.316700	0.134761	0.174324	0.208774	0.259249	0.293702	0.094481
24	[1.347, 1.407)	1.376290	0.140588	0.182656	0.218515	0.262003	0.296059	0.097091
25	[1.407, 1.468)	1.437290	0.139425	0.180701	0.218145	0.266967	0.305932	0.096661
26	[1.468, 1.528)	1.497320	0.143491	0.183022	0.220029	0.266648	0.313635	0.099192
27	[1.528, 1.588)	1.557600	0.141067	0.183644	0.222224	0.274634	0.306337	0.098174
28	[1.588, 1.649)	1.618450	0.146120	0.191419	0.232781	0.279342	0.307745	0.101508
29	[1.649, 1.709)	1.678350	0.148991	0.187539	0.221236	0.266810	0.290851	0.102900
30	[1.709, oo)	2.494600	0.180353	0.228782	0.271082	0.321884	0.359723	0.122037

Source: own

Note: The results were generated for $n = 4$ on the basis of the presented simulation framework. The outcome is based on 100,000 perturbed reciprocal PCMs. The simulation scenario assumed LLSM as the PT and Saaty's preference scale.

(and that difference is 0.62 - 0.24 = 0.38) then with 95% certainty it can be assumed that the highly ranked option is the most preferred option despite the judgment inconsistency and errors it entails for the vectors of priority ratios estimates.

However, assuming the same level of PCM consistency but a higher level of certainty e.g. 98%, a decision maker cannot assume that the first alternative will remain the highly ranked option despite the judgments inconsistency. It is so because the 0.98-quantile of the maximum absolute deviation distribution for $CI_{(LLSM)} = 0.109736$ equals 0.257923 (Tab. 3) and 0.98-quantile of the second maximal absolute deviation distribution equals 0.19094. In this case 0.19094 + 0.257923 = 0.448863 what is more than 0.38 and that means a decision maker cannot assume with 98% certainty that the two highly ranked alternatives cannot change their rankings.

Concluding, for the first time ever, it becomes possible to credibly and accurately relate a level of PCM consistency with priority vector credibility derived from inconsistent pairwise judgments, especially from the viewpoint of potential rank reversal of the first two highest ranked alternatives. For similar calculations but concerning a different number of alternatives, similar data can be elaborated and utilized. Exemplary data for three, five and six alternatives is presented in the appendix to this research paper.

3.3 Examination Breakthrough

Having in mind that a consistency index for the PCM denoting group preferences cannot be greater than the consistency index of the most inconsistent individual PCM (see e.g. Liu, Zhang, & Wang, 2012; Grošelj & Stirn, 2012; Escobar, Aguarón, & Moreno-Jimenez, 2004; Xu, 2000), it finally also becomes possible to designate priority vector credibility for the group on the basis of the most inconsistent individual PCM.

In this case it is simply suggested to take the most inconsistent individual PCM and designate on its basis (its consistency index) the possible priority ratios deviations. Then, the following steps are recommended: (a) calculate the PCM for the group on the basis of individual PCMs with application of the weighted geometric mean procedure presented earlier in this paper (individual judgments of decision makers are raised to their voting power and the group outcome is established on the basis of the weighted geometric mean for individual judgments); (b) compute a group priority vector, on the basis of the PCM established for the group, with application of e.g. LLSM; (c) apply earlier established priority ratios deviations of the most inconsistent individual PCM to the vector of priority ratios established for the group. In the case of the entire AHP framework, proceed similarly as during standard AHP weighing and adding procedure – weigh (by criteria and sub-criteria) and add either priority vectors and consistency indices of the most inconsistent individual PCMs and transfer designated deviations to the final priority vector in the way described above.

Now, why is the application of maximum absolute deviations among priority ratios suggested instead of average deviations? In the subsection entitled 'Examination methodology', it was emphasized that only a few papers dealt in depth with the problem concerning the relation between the level of pairwise judgments (in)consistency and the degree of possible estimation errors for established priority vector. As mentioned then, that research examined only average estimation errors within particular priority vectors which makes those results questionable, especially when relative measures for estimation errors are considered. These issues will be scrutinized briefly now.

The simulation framework presented in subsection 'Examination methodology' was applied to evaluate if average measures for estimation errors can be used. Instead of maximal absolute deviations calculated in Step 6, the average absolute deviations among absolute deviations and the mean absolute deviations – AD(MAD) (Formula 12), and the average absolute deviations among relative deviations and the mean relative deviations – AD(MRD) (Formula 13) were computed.

$$AD(MAD) = \frac{1}{n}\sum_{i=1}^{n}\left| |w_i - v_i| - \frac{1}{n}\sum_{i=1}^{n}|w_i - v_i| \right| \quad (12)$$

$$AD(MRD) = \frac{1}{n}\sum_{i=1}^{n}\left| \frac{|w_i - v_i|}{w_i} - \frac{1}{n}\sum_{i=1}^{n}\frac{|w_i - v_i|}{w_i} \right| \quad (13)$$

An examination of the above deviations designated by Formulae 12 and 13 within the

simulation framework presented in subsection 'Examination methodology' enables discovery whether single deviations between estimated and 'true' specific priority ratios are close to the average deviation between estimated and 'true' vector of priority ratios. If that was the case, both examined measures i.e. *AD(MAD)* and *AD(MRD)* should be close to zero and their performance should not depend on values of the examined consistency index. Fig. 3 and 4, which present selected relations between *AD(MAD)* or *AD(MRD)* and $CI_{(LLSM)}$ clearly contradict the above assumption about single deviations between estimated and 'true' specific priority ratios i.e. single deviations are not close to the average deviation between estimated and 'true' priority vector but they fluctuate.

It means that one can expect strong variability among single priority ratios deviations within a given priority vector. Thus, a mean deviation between the 'true' and the estimated priority vector should not be considered as a good indicator of possible estimation errors, because the single deviations are not close

to the mean deviation calculated for the entire priority vector.

To illustrate, the following hypothetic situation is presented as an example. Let PV(*w*) = [0.2262, 0.2729, 0.1390, 0.3619] denotes the 'true' vector of the priority ratios and PV(*v*) = [0.2143, 0.3571, 0.1429, 0.2857] denotes the estimated vector of priority ratios. Mean relative and absolute deviations for those vectors equal *MRD* = 0.15 and *MAD* = 0.0441, respectively. At the same time single relative and absolute deviations among specific priority ratios fluctuate from 0.028 to 0.3087 in the case of *MRD*, and from 0.0039 to 0.0842 in the case of *MAD*. Thus, taking the *MRD* as the example, one could expect the average relative deviation at the level of 15% while its single value can fluctuate from around 3% to around 31%. This is why mean deviations (especially relative ones) are not good indicators of possible estimation errors for priority vectors and should not be taken into consideration during further research.

Fig. 3:   Performance of $CI_{(LLSM)}$ in relation to *AD(MAD)* and *AD(MRD)*
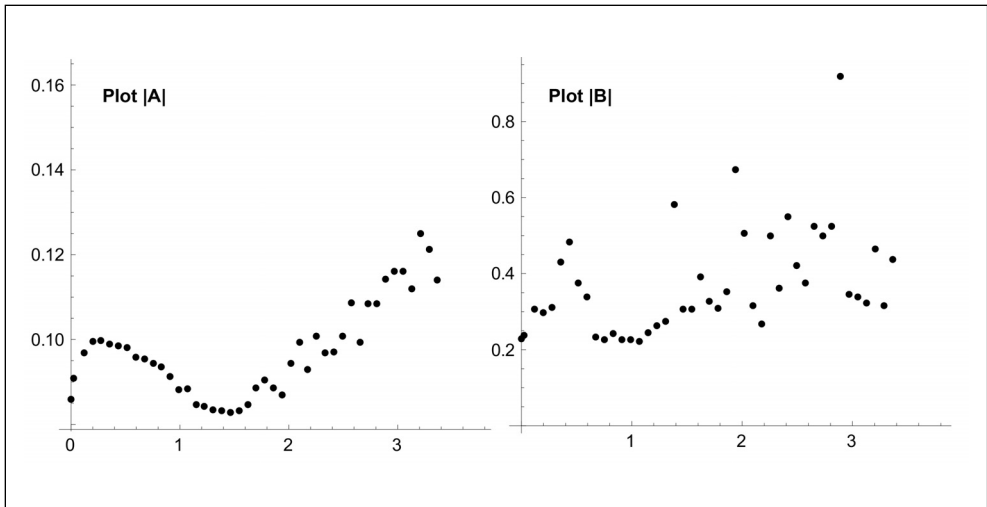
Source: own

Note: Plots are based on 18,000 random reciprocal PCMs for *n* = 3. The results are generated with application of LLSM as the PT and Saaty's preference scale. Plot |A| presents the relation among $CI_{(LLSM)}$ and 0.95-quantile of *AD(MAD)*. Plot |B| presents the relation between $CI_{(LLSM)}$ and a mean *AD(MRD)*.

Fig. 4: Performance of $CI_{(LLSM)}$ in relation to $AD(MAD)$ and $AD(MRD)$
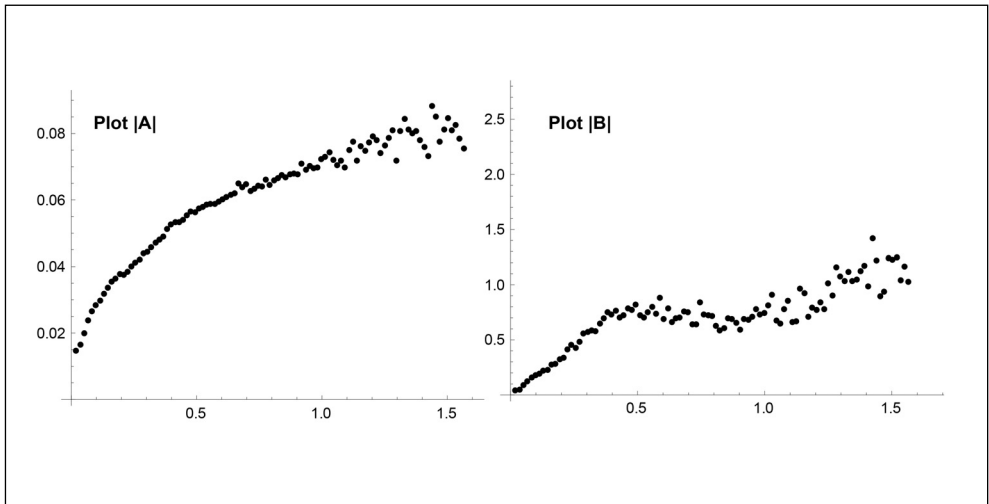
Source: own

Note: Plots are based on 18.000 random reciprocal PCMs for $n = 6$. The results are generated with application of LLSM as the PT and Saaty's preference scale. Plot |A| presents the relation between $CI_{(LLSM)}$ and 0.95-quantile of $AD(MAD)$. Plot |B| presents the relation among $CI_{(LLSM)}$ and a mean $AD(MRD)$.

Conclusions

The scope of the research concerns issues associated with group decision making (GDM) as the most challenging process which entails various viewpoints and preferences of individuals that must be taken into consideration and somehow combined into one meaningful outcome.

When group decision making process was taken into consideration, the AHP seemed a particularly attractive methodology to examine the process. Although the methodology has been evaluated numerous times from the perspective of its effectiveness and applicability in GDM processes, a research gap has been identified and examined in this research paper. The impact of which inconsistency of judgments makes on confidence about priority ratios (PRs) of the priority vector (PV) established for a group of decision makers (DMs) has been studied. Although individual relations have been studied, examination results generalize to the synthesized pairwise comparison matrix (SPCM) that is obtained on the basis of individual pairwise comparison matrices for all group members.

The relation between the consistency of PCM and the PV estimate credibility has been thoroughly examined. The examination process has proceeded with the application of Monte Carlo simulations coded and executed in Wolfram Mathematica Software. To properly examine the problem, a sophisticated simulation algorithm has been elaborated and applied.

Having in mind that a consistency index for the PCM denoting group preferences cannot be greater than the consistency index of the most inconsistent individual PCM, it became possible to designate the credibility of the preference vector for the group on the basis of the most inconsistent individual PCM. The examination's conclusion is based on the fundamental assumption that individual judgments of decision makers are raised to their voting power and the group outcome is established on the basis of their individual judgments i.e. the weighted geometric mean is formed. Then, and only then, the credibility of the group outcome can be designated with the application of the consistency index of the most inconsistent individual PCM.

It is emphasized that thus far only few papers have dealt with the problem concerning the relation between the level of the pairwise judgments (in)consistency and the size of possible estimation errors for established vector of priority ratios. However, the outcome of those papers is limited because those research efforts examined only average estimation errors within particular priority vectors which made those results less accurate as compared to this research paper, especially when relative measures for estimation errors were applied. This research paper overcomes limitations of those other examinations which distinguishes it from other papers and emphasizes its novelty.

Certainly, there is a place for further examinations in the research area e.g. other consistency indices and other prioritization techniques can be studied from the perspective of this research effort's objectives to enhance its findings and broaden its methodological implications.

References

Abdelmaguid, T. F., & Elrashidy, W. (2016). Halting decisions for gas pipeline construction projects using AHP: a case study. *Operational Research, 19*(1), 1–21. https://doi.org/10.1007/s12351-016-0277-2.

Aczél, J., & Saaty, T. L. (1983). Procedures for synthesizing ratio judgments. *Journal of Mathematical Psychology, 27*(1), 93–102. https://doi.org/10.1016/0022-2496(83)90028-7.

Aguarón, J., & Moreno-Jimenez, J. M. (2003). The geometric consistency index: Approximated thresholds. *European Journal of Operational Research, 147*(1), 137–145. http://dx.doi.org/10.1016/S0377-2217(02)00255-2.

Altuzarra, A., Moreno-Jiménez, J. M., & Salvador, M. (2010). Consensus building in AHP-group decision making: a Bayesian approach. *Operations Research, 58*(6), 1755–1773.

Aguarón, J., Escobar, M. T., & Moreno-Jiménez, J. M. (2014). The precise consistency consensus matrix in a local AHP-group decision making context. *Annals of Operations Research, 245*(1-2), 1–15. http://dx.doi.org/10.1007/s10479-014-1576-8.

Bana e Costa, C. A., & 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. http://dx.doi.org/10.1016/j.ejor.2006.09.022.

Barzilai, J. (2005). Measurement and preference function modeling. *International Transactions in Operational Research, 12*(2), 173–183. http://doi.org/10.1111/j.1475-3995.2005.00496.x.

Belton, V., & Gear, T. (1983). On a shortcoming of Saaty's method of analytic hierarchies. *Omega, 11*(3), 228–230.

Bozóki, S., & Rapcsák, T. (2008). On Saaty's and Koczkodaj's inconsistencies of pairwise comparison matrices. *Journal of Global Optimization, 42*(2), 157–175. https://doi.org/10.1007/s10898-007-9236-z.

Brunelli, M. (2018). A survey of inconsistency indices for pairwise comparisons. *International Journal of General Systems, 47*(8), 751–771. https://doi.org/10.1080/03081079.2018.1523156.

Budescu, D. V., Zwick, R., & Rapoport, A. (1986). A comparison of the eigenvalue method and the geometric mean procedure for ratio scaling. *Applied Psychological Measurement, 10*(1), 69–78.

Condorcet, M. J. A., de (1785). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Paris: De l'Imprimerie Royale [facsimile edition New York, NY: Chelsea, 1972].

Crawford, G., & Williams, C. (1980). *Analysis of subjective judgment matrices*. Santa Monica, CA: Rand Corporation R-2572-AF.

Crawford, G., & Williams, C. (1985). A note on the analysis of subjective judgment matrices. *Journal of Mathematical Psychology, 29*(4), 387–405.

Dijkstra, T. K. (2013). On the extraction of weights from pairwise comparison matrices. *Central European Journal of Operations Research, 21*(1), 103–123. http://dx.doi.org/10.1007/s10100-011-0212-9.

Dixit, P. D. (2018). Entropy production rate as a criterion for inconsistency in decision theory. *Journal of Statistical Mechanics: Theory and Experiment,* (5), 353–408.

Dong, Y., Xu, Y., Li, H., & Dai, M. (2008). A comparative study of the numerical scales and the prioritization methods in AHP. *European Journal of Operational Research, 186*(1), 229–242. https://doi.org/10.1016/j.ejor.2007.01.044.

Escobar, M. T., Aguarón, J., & Moreno-Jiménez, J. M. (2004). A note on AHP group consistency for the row geometric mean

prioritization procedure. *European Journal of Operational Research, 153*(2), 318–322. http://doi.org/10.1016/S0377-2217(03)00154-1.

Fedrizzi, M., & Ferrari, F. (2017). A chi-square-based inconsistency index for pairwise comparison matrices. *Journal of the Operational Research Society, 69*(7), 1125–1134. https://doi.org/10.1080/01605682.2017.1390523.

Grošelj, P., & Stirn, L. Z. (2012). Acceptable consistency of aggregated comparison matrices in Analytic Hierarchy Process. *European Journal of Operational Research, 223*(2), 417–420. https://doi.org/10.1016/j.ejor.2012.06.016.

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. http://dx.doi.org/10.1016/j.eswa.2015.08.049.

Grzybowski, A. Z. (2012). Note on a new optimization based approach for estimating priority weights and related consistency index. *Expert Systems with Applications, 39,* 11699–11708. http://dx.doi.org/10.1016/j.eswa.2012.04.051.

Hosseinian, S. S., Navidi, H., & Hajfathaliha, A. (2012). A new linear programming method for weights generation and group decision making in the Analytic Hierarchy Process. *Group Decision and Negotiation, 21*(3), 233–254. https://doi.org/10.1007/s10726-009-9182-x.

Johnson, C. R., Beine, W. B., & Wang, T. J. (1979). Right-left asymmetry in an eigenvector ranking procedure. *Journal of Mathematical Psychology, 19*(1), 61–64.

Kazibudzki, P. T. (2016b). An examination of performance relations among selected consistency measures for simulated pairwise judgments. *Annals of Operations Research, 244*(2), 525–544. http://dx.doi.org/10.1007/s10479-016-2131-6.

Kazibudzki, P. T. (2019a). An examination of ranking quality for simulated pairwise judgments in relation to performance of the selected consistency measure. *Advances in Operations Research*, *2019,* 1–24. https://doi.org/10.1155/2019/3574263.

Kazibudzki, P. T. (2019b). The quality of ranking during simulated pairwise judgments for examined approximation procedures. *Modelling and Simulation in Engineering, 2019,* 1–13. https://doi.org/10.1155/2019/1683143.

Koczkodaj, W. W., & Urban, R. (2018). Axiomatization of inconsistency indicators for pairwise comparisons. *International Journal of Approximate Reasoning, 94,* 18–29. https://doi.org/10.1016/j.ijar.2017.12.001.

Koczkodaj, W. W., & Szwarc, R. (2014). On axiomatization of inconsistency indicators for pairwise comparisons. *Fundamenta Informaticae, 132*(4), 485–500. http://doi.org/10.3233/FI-2014-1055.

Kou, G., Ergu, D., Lin, C., & Chen, Y. (2016). Pairwise comparison matrix in multiple criteria decision making. *Technological and Economic Development of Economy, 22*(5), 738–765. http://doi.org/10.3846/20294913.2016.1210694.

Kramulová, J., & Jablonský, J. (2016). AHP model for competitiveness analysis of selected countries. *Central European Journal of Operations Research, 24*(2), 335–351. https://doi.org/10.1007/s10100-015-0394-7.

Kułakowski, K. (2015). A heuristic rating estimation algorithm for the pairwise comparisons method. *Central European Journal of Operations Research, 23*(1), 187–203. https://doi.org/10.1007/s10100-013-0311-x.

Lidinska, L., & Jablonsky, J. (2018). AHP model for performance evaluation of employees in a Czech management consulting company. *Central European Journal of Operations Research, 26*(1), 239–258. https://doi.org/10.1007/s10100-017-0486-7.

Linares, P. (2009). Are Inconsistent Decisions Better? An Experiment with Pairwise Comparisons, *European Journal of Operational Research, 193*(2), 492–498.

Linares, P., Lumbreras, S., Santamaría, A., & Veiga, A. (2016). How relevant is the lack of reciprocity in pairwise comparisons? An experiment with AHP. *Annals of Operations Research, 245*(1-2), 227–244. https://doi.org/10.1007/s10479-014-1767-3.

Liu, F., Zhang, W.-G., & Wang, Z.-X. (2012). A goal programming model for incomplete interval multiplicative preference relations and its application in group decision making. *European Journal of Operational Research, 218*(3), 747–754. https://doi.org/10.1016/j.ejor.2011.11.042

Merkin, B. G. (1979). *Group choice*. New York, NY: John Wiley & Sons.

Mizuno, T. (2019). *A link diagram for pairwise comparisons*. In: Czarnowski I, Howlett RJ, Jain LC, and Vlacic L. (ed) (2018) *Intelligent Decision Technologies*. 181–186. Berlin: Springer International Publishing.

Moreno-Jiménez, J. M., Aguarón, J., & Escobar, M. T. (2008). The core of consistency

in AHP-group decision making. *Group Decision and Negotiation, 17*(3), 249–265. https://doi.org/10.1007/s10726-007-9072-z.

Moreno-Jimenez, J. M., Joven, J. A., Pirla, A. R., & Lanuza, A. T. (2005). A spreadsheet module for consistent consensus building in AHP-group decision making. *Group Decision and Negotiation, 14*(2), 89–108. https://doi.org/10.1007/s10726-005-2407-8.

Orbán-Mihálykó, É., Mihálykó, C., & Koltay, L. (2017). A generalization of the Thurstone method for multiple choice and incomplete paired comparisons. *Central European Journal of Operations Research, 27*(1)*,* 133–159. https://doi.org/10.1007/s10100-017-0495-6.

Peláez, J. I., Martínez, E. A., & Vargas, L. G. (2018). Consistency in positive reciprocal matrices: an improvement in measurement methods. *IEEE access 6*: 25600–25609.

Ponis, S. T., Gayialis, S. P., Tatsiopoulos, I. P., Panayiotou, N. A., Stamatiou, D.-R. I., & Ntalla, A. C. (2015). An application of AHP in the development process of a supply chain reference model focusing on demand variability. *Operational Research, 15*(3), 337–357. https://doi.org/10.1007/s12351-014-0163-8.

Saaty, T. L. (1980). *The Analytic Hierarchy Process*. New York, NY: McGraw Hill.

Saaty, T. L. (2006). *Fundamentals of decision making and priority theory with the Analytic Hierarchy Process*. Pittsburgh, PA: RWS Publication.

Saaty, T. L. (2008a). Decision making with the Analytic Hierarchy Process. *International Journal of Services Sciences, 1*(1), 83–98. https://doi.org/10.1504/IJSSCI.2008.017590.

Saaty, T. L. (2008b). Relative measurement and its generalization in decision making. Why pairwise comparisons are central in mathematics for the measurement of intangible factors. The Analytic Hierarchy/Network Process. *Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales. Serie A, Matemáticas, 102*(2), 251–318. https://doi.org/10.1007/BF03191825.

Saaty, T. L., & Peniwati, K. (2008). *Group decision making*. Pittsburgh, PA: RWS Publications.

Saaty, T. L., & Vargas, L. G. (1984). Comparison of eigenvalue, logarithmic least square and least square methods in estimating ratio. *Mathematical Modeling, 5*(5), 309–324. https://doi.org/10.1016/0270-0255(84)90008-3.

Saaty, T. L., & Vargas, L. G. (2012). The possibility of group choice: pairwise comparisons and merging functions. *Social Choice and Welfare, 38*(3), 481–496.

Scala, N. M., Rajgopal, J., Vargas, L. G., & Needy, K. L. (2016). Group decision making with dispersion in the Analytic Hierarchy Process. *Group Decision and Negotiation, 25*(2), 355–372. https://doi.org/10.1007/s10726-015-9445-7.

Schoner, B., & Wedley, W. C. (1989). Ambiguous criteria weights in AHP: Consequences and solutions. *Decision Sciences, 20*(3), 462–475. https://doi.org/10.1111/j.1540-5915.1989.tb01561.x.

Sun, L., & Greenberg, B. S. (2006). Multicriteria Group Decision Making: Optimal Priority Synthesis from Pairwise Comparisons. *Journal of Optimization Theory and Applications, 130*(2), 317–338.

Temesi, J. (2011). Pairwise comparison matrices and the error-free property of the decision maker. *Central European Journal of Operations Research, 19*(2), 239–249. https://doi.org/10.1007/s10100-010-0145-8.

Thurstone, L. L. (1927). A law of comparative judgments. *Psychological Reviews, 34,* 273–286.

Wu, W., & Kou, G. (2016). A group consensus model for evaluating real estate investments alternatives. *Financial Innovation, 2*(8), 1–10. https://doi.org/10.1186/s40854-016-0027-8.

Xu, W. J., Dong, Y. C., & Xiao, W. L. (2008). Is it reasonable for Saaty's consistency test in the pairwise comparison method? *Proceedings of 2008 ISECS International Colloquium on Computing, Communication, Control, and Management, 3,* 294–298.

Xu, Z. S. (2000). On consistency of weighted geometric mean complex judgment matrix in AHP. *European Journal of Operations Research, 126*(3), 683–687. https://doi.org/10.1016/S0377-2217(99)00082-X.

Young, H. P. (1988). Condorcet's theory of voting. *American Political Science Review, 82*(4), 1231–1244. https://doi.org/10.2307/1961757.

Zahedi, F. (1986). A simulation study of estimation methods in the analytic hierarchy process. *Socio-Economic Planning Science, 20*(6), 347–354. https://doi.org/10.1016/0038-0121(86)90046-7.

Ing. Pawel Tadeusz Kazibudzki, PhD.
Opole University of Technology
Faculty of Economics and Management
Department of Enterprise Organization
and Management
Poland
p.kazibudzki@po.edu.pl

doc. Ing. Jiří Křupka, PhD.
University of Pardubice
Faculty of Economics and Administration
Institute of Engineering Systems
and Informatics
Czech Republic
jiri.krupka@upce.cz

Abstract

PAIRWISE JUDGMENTS CONSISTENCY IMPACT ON QUALITY OF MULTI-CRITERIA GROUP DECISION-MAKING WITH AHP

Pawel Tadeusz Kazibudzki, Jiří Křupka

The scope of this research encompasses issues associated with group decision making (GDM) as the most challenging process which entails various viewpoints and preferences of individuals that must be taken into consideration and somehow combined into one meaningful outcome. When GDM is taken into consideration, the AHP seems to be a particularly attractive methodology. From the perspective of its applications, an existing research gap has been identified and examined in this research paper. Thus, the inconsistency of judgments impact on priority vector quality has been examined from the perspective of group decision making. Examination results generalize to the synthesized pairwise comparison matrix that is obtained on the basis of individual pairwise comparison matrices for all group members. The examination process has proceeded with the application of Monte Carlo simulations coded and executed in Wolfram Mathematica Software. Having in mind that a consistency index for the PCM denoting group preferences cannot be greater than the consistency index of the most inconsistent individual PCM it became possible to designate the credibility of the priority vector for the group on the basis of the most inconsistent individual PCM. It is emphasized that thus far only a few papers have dealt with the problem concerning the relation between a level of the pairwise judgments inconsistency and the degree of possible estimation errors for established vector of priority ratios. This research paper overcomes limitations of other examinations which distinguishes it from other papers and emphasizes its novelty.