Analysis of Errors in Priority Vector Estimation and Their Relationship with the Correctness of the Final Ranking of Decision Alternatives

Andrzej Z. Grzybowski, *Member, IAENG,* Tomasz Starczewski

Abstract—**Ranking creation is the main purpose of the multiple-criteria decision analysis (MCDA). In practice, it is typically achieved by estimation of the priority weights that reflect the importance of each of the available alternatives - this process is called prioritization. One of the most popular MCDA methodologies is the Analytic Hierarchy Process (AHP). All priority-weights-estimation technics that are used under the AHP scheme are based on the so-called pairwise comparison matrix (PCM). The PCM elements represent the decision maker's judgments about the priority-weights-ratios. They are typically expressed in values from a predefined set of numbers that is called a scale. Because of human brain limitations, it is natural that these judgments are usually erroneous, and consequently, the estimates of the priority weights are erroneous as well. It is well understood, however, that serious errors in judgments make the information contained in the PCM worthless. Thus the decision makers need a procedure that enables them to accept a given PCM or reject it as a useless one. This paper is devoted to the simulation analysis of the prioritization errors and their relationship with the correctness of the final ranking of decision-alternatives. Our simulation experiments reveal some interesting facts about the impact of the adopted scale on the priority-weights-estimation errors and allow us to formulate important remarks about the PCM acceptance procedure.**

Index Terms—**AHP, prioritization, estimation errors, final ranking, simulation.**

I. Introduction

MULtiple-criteria decision analysis (MCDA) is a branch of multiple criteria decision making that deals with problems that have only a small number of alternatives that have to be ranked. The applications of MCDA cover very wide range of real-world problems including such different tasks as supplier selection [5], energy selection [15], comparison of bridge designs [9], evaluation of e-commerce websites qualities [3] or cloud service provider selection [27] - to name just a few of many interesting problems described in literature. In order to solve these problems the decision maker needs to create ranking of the available decision alternatives. In the MCDA practice, the alternatives' ranking creation is typically achieved by estimation of so-called priority weights, i.e. numbers telling to what degree a given alternative satisfies a given criterion. Apart from the alternatives and a number of criteria, more complex MCDA problems may also involve several experts and/or decision makers (DM). To obtain the final ranking of the available alternatives all these factors need to be ranked as well. It is confirmed in recent literature, [29], that one of the most frequently used prioritization methodologies is the Analytic Hierarchy Process (AHP), [22]. In the AHP all problem's factors are arranged in a hierarchical structure descending from - say - decision makers to criteria and decision-alternatives in successive levels. After all sub-problems of prioritization are solved (i.e. the rankings of DMs, experts, criteria and alternatives are obtained), the final ranking of decision-alternatives' is created. All priority-weights-estimation technics that are used under the AHP scheme are based on the so-called pairwise comparison matrix (PCM). The PCM contains the decision maker's judgments about the priority-weights-ratios. These judgments are typically expressed in values from an adopted predefined set of numbers that is called a scale, [21], [22], [4], [8], [10], [19], [26]. However, because of human brain natural limitations, in praxis, the DM's judgments are usually erroneous and as a consequence, the estimates of the priority weights derived on their basis are erroneous as well. Thus, in a sense, it is natural and unavoidable to accept some level of judgments' incorrectness. However, on the other hand, it is well-understood that serious errors in the DM's judgments make the information contained in the PCM worthless. Thus an important problem within the AHP methodology is the ability to distinguish between useful PCMs and the useless ones.

This paper is devoted to the simulation analysis of the estimation errors and their relationship with the correctness of the final ranking of alternatives. Its gist follows the concept described in [14] and extends results presented therein. Section 2 introduces all necessary notions and definitions. In Section 3 the considered problem is described in details. Section 4 presents adopted simulation frameworks and discusses the results of our simulation experiments. This article is concluded with final remarks about the impact of the adopted scale as well as about the foundations of a new rational PCM acceptance procedure.

II. Preliminaries: notation and basic facts

A priority vector (PV) is a vector of priority weights $v = (v_1, \ldots, v_n)'$, i.e. numbers that reflect the absolute degree/intensity of importance of each alternative with respect to a given criterion. A basic assumption of the AHP is that for each criterion considered in a given MCDA problem, the PV does exist and that it is unique up to a multiplying constant. Typically, the priority weights $v_i,\ i = 1, \ldots, n$, are positive and the PV is normalized to unity, [24].

Another basic assumption of the AHP is that the decision-maker is able to evaluate the ratios of priorities $a_{ij} = v_i/v_j$.

Manuscript received June 11, 2019; revised July 11, 2019

A.Z. Grzybowski and T. Starczewski are with the Institue of Mathematics, Faculty of Mechanical Engineering and Computer Science, Czestochowa University of Technology, Czestochowa, 42-201 Poland, e-mail: andrzej.grzybowski@im.pcz.pl, tomasz.starczewski@im.pcz.pl

As a result of such comparisons, the pairwise comparison matrix $\mathbf{A} = [a_{ij}]_{nxn}$ is obtained. In the conventional AHP, the input data of the PCM is collected only for the upper triangle of the matrix $\mathbf{A}$, while the remaining elements are computed as the inverses of the corresponding symmetric elements in the upper triangle i.e. $a_{ij} = 1/a_{ji}$ for $i > j$. A PCM that satisfies the latter condition is said to be *reciprocal*. A PCM is said to be a *consistent* one, if it is reciprocal and its elements satisfy the condition:

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

As we have mentioned above, it is obvious that in practice one cannot expect that the elements of PCM give the priority ratios precisely. First of all, according to the usual procedure, the DM's answers are given in lexical phrases and then they are transformed into numbers belonging to a given scale. Usually these scales are sets of up to a few dozens of numbers. In such a case one cannot neglect rounding errors. The most popular scale is the Saaty's one (SS). The SS contains integers 1,2,...,9 and their reciprocals. Other scales suggested in literature are the Extended Saaty's scale (ESS(N)) that contains integers from 1 to N, along with their reciprocals, and the geometric scale GS(c) that contains numbers s of the form $s = c^{i/2},\ i \in I$ with I being a predefined set of integers. A deeper description of various AHP scales can be found e.g. in [8], [26].

Apart from the rounding errors, there are also other kinds of errors in the ratios-evaluations that are results of human brain limitations. Consequently, even if the comparisons are done very carefully, PCM in reality is inconsistent. Because serious errors in DM's judgments can result in a misleading PCM, an important problem connected with this methodology is how to measure the degree of inconsistence of the PCM. But before we briefly introduce some inconsistency indices, first we need to recall two most popular prioritization methods. One of them is the so-called logarithmic least squares method, also known as geometric mean method - GM, [7]. The estimated priority vector (EPV) in the GM can be obtained by the following formula:

$$w_i = \left(\prod_{i=1}^{n} a_{ij}\right)^{1/n} \Bigg/ \sum_{i=1}^{n} \left(\prod_{i=1}^{n} a_{ij}\right)^{1/n}$$

Another, perhaps the most commonly used prioritization method is based on specific results from the matrix spectral theory. This one is called right eigenvector method (REV). Its description and formal backgrounds have vast literature, see e.g. [22], [23] and many contemporary books devoted to the AHP. Thus the details of the underlying reasoning are omitted here to save the article space. In short, the REV method suggests to take as the EPV, the eigenvector associated with the principal eigenvalue of the PCM at hand. To achieve this goal, one needs to solve the characteristic equation

$$det(\mathbf{A} - \lambda \mathbf{I}) = 0$$

and then the linear equation (with respect to $\mathbf{w}$)

$$\mathbf{A}\mathbf{w} = \lambda_{\max}\mathbf{w}$$

where $\lambda_{\max}$ is the principal eigenvalue of $\mathbf{A}$ (i.e. the largest solution of the previous characteristic equation). The existence and uniqueness of the principal eigenvector $\mathbf{w}$ is guaranteed by the Perron's theorem, that also assures the positivity of all its coordinates.

There are many papers that present results of studies of the performance of these and other prioritization methods. In these studies it is found, e.g. [11], [6], [20], that the EPVs obtained with the help of the GM and REV differ very little. There is also no agreement which one is better - both of the methods have their pros and cons. However, the EPV is certainly much more easy to compute via the GM, thus this method will be primarily used in our studies.

As we indicated, apart from deriving priority vectors, another problem within the AHP methodology is how to measure the degree of inconsistency of the PCM. We are presented with a number of inconsistency indices and again, the two most frequently used ones are related to the two above introduced prioritization methods.

The index connected with the REV, denoted as SI, was proposed by Saaty and is defined as follows:

$$SI(n) = \frac{\lambda_{\max} - n}{n - 1}$$

Related to the geometric mean method index (GI) was proposed by Crawford and Williams[7], and was popularized by Aguaron and Moreno-Jimenez [1] in 2003. It is given by

$$GI(n) = \frac{2}{(n-1)(n-2)} \sum_{i<j} \log^2(a_{ij}w_i/w_j)$$

Apart from these two indices there is third popular index that is based on the notion of the triad inconsistency. It was proposed by Koczkodaj [18]. Following him, for any different $i, j, k \leq n$, a tuple (a_{ik}, a_{ij}, a_{kj}) will be called a *triad*. Koczkodaj proposed to characterize the triad's inconsistency by the number:

$$TI(a_{ik}, a_{kj}, a_{ij}) = \min\left[\left|1 - a_{ik}a_{kj}/a_{ij}\right|, \left|1 - a_{ij}/((a_{ik}a_{kj})\right|\right]$$

Then, the Koczkodaj's inconsistency index KI of any reciprocal PCM is defined as a maximum of triad's inconsistencies i.e. $KI = max[TI(a_{ik}, a_{kj}, a_{ij})]$, where the maximum is taken over all triads in the upper triangle of the PCM.

Yet another inconsistency index that manifests very good correlation with PV estimation quality was defined as the average value of all triad's inconsistencies. It was introduced and studied in [12] and is denoted here as ATI.

All inconsistency indices are developed in order to enable the DM to distinguish between useful and useless PCMs. Thus, usually all these indices are given along with related consistency thresholds. However, as it was criticized by many researchers, typically the values of acceptance thresholds are based on some heuristics and are not supported by any profound formal reasoning or statistical research. In recent literature there are many confusing examples that prove that the thresholds work poorly, see [11] and literature in there. To deal with this problem a new approach to PCM acceptance was proposed quite recently in [12], where the relationship between the inconsistency indices and the magnitude of priority-vector-estimation-error (PVEE) was examined with the help of Monte Carlo simulations. In [12] it was proposed to measure the PVEE as the average absolute (AE) and/or relative (RE) errors. The errors are given by the following formulae:

$$AE(\mathbf{v}, \mathbf{w}) = \frac{1}{n}\sum_{i=1}^{n} |v_i - w_i| \quad (1)$$

$$RE(\mathbf{v}, \mathbf{w}) = \frac{1}{n}\sum_{i=1}^{n} \frac{|v_i - w_i|}{v_i} \quad (2)$$

where $\mathbf{v} = (v_1, ..., v_n)$ is the true PV while $\mathbf{w} = (w_1, ..., w_n)$ is its EPV. Obviously the EPV and consequently the errors depend on the prioritization method as well. So, in our studies both the GM and REV were used for PV estimation and then calculation of AE and RE. However the numeric results presented here are related to the GM.

III. Problem Statement

As we have already indicated, it is argued that the knowledge about the relationship between inconsistency indices and PVEEs would help the DM in making decisions about the PCM acceptance because such decisions are based on the observed value of the inconsistency index. The principal goal of this paper is to help the DM in such a task. To achieve this goal one should be able to distinguish between ”small”, ”average” and ”big” PVEEs. However, as yet there are no criteria for making such a ”classification” of the estimation errors. To cope with this problem we proposed in [14], for the first time in literature, to investigate the relationship between the PVEEs and the chances of ”significantly wrong” final EPV. The latter term as well as the idea of the proposed criterion needs additional clarification. Below we present more formal and more detailed description of our proposal.

During the analysis of the multi-criteria- decision-making problem under the AHP scheme, both the prioritization methods and the inconsistency analysis are curried out several times (as described in Introduction, at least for the criteria and then for the alternatives with respect to each criterion separately). By PCM(Cr) we denote the PCM that was provided by the decision-maker for the criteria, and by the symbol PCM(i) we denote the PCMs achieved for the decision alternatives with respect to the i-th criterion. Obviously, the PCM(Cr) has order k, where k is the number of different criteria in the considered decision-making problem, $(k > 1)$. Similarly, the PCM(i) has the order n, with n being, as previously, the number of available decision alternatives. Let $\mathbf{v}^0$ and $\mathbf{w}^0$ be, respectively, the true PV for criteria and the EPV computed for the criteria on the basis of PCM(Cr). Let also $\mathbf{v}^i$ and $\mathbf{w}^i$ be the true PV and its EPV for the alternatives with respect to the i-th criterion. The true final ranking of the alternatives is given by the final true PV - say $\mathbf{v}$ - that is equal, by the definition, to the weighted average of the vectors $\mathbf{v}^i,\ i = 1, \ldots, k$, with the weights given in $\mathbf{v}^0 = (v_1^0, v_2^0, \ldots, v_k^0)$, i.e.

$$\mathbf{v} = \sum_{i=1}^{k} v_i^0 \mathbf{v}^i \quad (3)$$

Analogously, the estimated final PV - say $\mathbf{w}$ - is equal to the weighted average of the vectors $\mathbf{w}^i,\ i = 1, , k$, with the weights given in $\mathbf{w}^0 = (w_1^0, w_2^0, \ldots, w_k^0)$

Now we introduce a crucial notion for our research. We will say, that $\mathbf{w}$ provides us with *significantly incorrect final ranking* if the truly best alternative (i.e. the one associated with the greatest coordinate in $\mathbf{v}$) is not the best one in the estimated final ranking (i.e. its corresponding component in $\mathbf{w}$ is not the greatest one) Hereafter we will abbreviate the phrase ”significantly incorrect final ranking” to SIFR.

We propose such a definition to focus only on serious mistakes in rankings, so our criterion does not take into account such situations where the final classification is wrong, but the erroneous ranks are related to less significant alternatives. That is why we call it ”significantly wrong”.

From the multiple-criteria decision analysis point of view, not the PVEEs themselves but the chances for significantly incorrect final ranking is perhaps the most important criterion for the PCM acceptance or rejection. After all, if the final ranking is correct, the magnitude of errors in estimates of weights are not that meaningful. So, the first aim of our studies is to find out what is the relationship between the magnitude of the PVEEs and the probabilities of obtaining SIFR.

Another interesting question is what type of an error is more meaningful, the AE or RE? In [12] it was suggested that in the context of the AHP applications, the more important one is the RE. However, this statement was supported only by intuition and some heuristics. Now we want to study which of the two types of errors are closer connected with the probability of obtaining the SIFR.

Yet another problem considered in our studies is the classification of the PVEEs, as mentioned at the beginning of this section. We focus here on the class of ”small” errors. In our opinion, as such can be obviously treated these errors that result from the methodology itself, that is the rounding errors. We study whether the magnitude of such ”small” errors depends on the adopted scale and - if so - which scale is the best one with respect to the criterion of the possibility of obtaining the SIFR.

All these questions can be answered only with the help of Monte Carlo simulation. Next section provides us with the description of the adopted simulation frameworks that are used in our studies.

IV. Simulation frameworks and results

To analyze the relationship between the PVEEs and the probabilities of obtaining the SIFR we use a simulation framework, that is a modification of the one adopted in [14].The main difference between these two frameworks is that now, instead of disturbing directly the ”true” PVs, first, we obtain the ”true” PCMs, and then we randomly disturb their elements. Next we round the disturbed elements to the nearest value from the adopted scale. Such a modification makes the simulations better imitate real phenomena that occur during such decision-making problems. Our framework for simulation experiments comprises the following steps.

Step 0 (Initialization) Set: n - the number of alternatives, k- the number of criteria, N -the number of simulated AHP problems, PR - the probability distribution of random PCM-estimation errors

Step 1 Randomly generate the ”true” priority vectors $\mathbf{v}^i,\ i = 0, , k$,

Step 2 On the basis of the vectors $\mathbf{v}^i,\ i = 0, \ldots, k$, compute related ”true” comparison matrix $\mathbf{M}^i$ with elements $m_{j,l}^i = \frac{v_j^i}{v_l^i}$

Step 3 Simulate "judgment errors" by disturbing the elements of PCM($\mathbf{v}^i$), $i = 0, \ldots, k$ according to the probability distribution PR.
Step 4 Compute rounded matrices $\mathbf{RM}^i, i = 0, ..., k$, by replacing the disturbed elements from the upper-triangle of $\mathbf{M}^i, i = 0, ..., k$, with the closest values from the adopted scale. Next replace all elements in the lower triangle of the $\mathbf{RM^i}$ with the reciprocities of the appropriate elements from the upper triangle.
Step 5 On the basis of the disturbed PCM($\mathbf{v}^i$), compute the "estimated" priority vectors $\mathbf{w}^i,\ i = 0, ..., k$, with the help of adopted prioritization method.
Step 6 With the help of formulae 3 compute the final "true" and "estimated" PVs: $\mathbf{v}$ and $\mathbf{w}$, respectively
Step 7 With the help of formulae 1 and 2 compute the PVEEs: $A_i = AE(\mathbf{v}^i, \mathbf{w}^i)$ and $R_i = RE(\mathbf{v}^i, \mathbf{w}^i)$, $i = 0, \ldots, k$.
Step 8 Compute the aggregated estimation errors as: $AAE = 1/(k+1) \sum_{i=0}^{k} A_i$ and $ARE = 1/(k+1) \sum_{i=0}^{k} R_i$
Step 9 Set $sifr = 1$ if the true final best alternative is different from the estimated final best alternative, otherwise set $sifr = 0$
Step 10 Write down all values computed and/or set in Steps 7 to 9 as one record.
Step 11 N times repeat Steps 1 to 10
Step 12 Return all records organized as one database.

As a result of simulation experiments conducted within the above simulation framework we receive a database that enables us to study the relationship between the PVEEs and the probabilities of obtaining the SIFR. For this purpose the whole database is arranged in ascending order with respect to the values of a considered type of errors $(A_i, R_i, AAE$ or $ARE)$ and then split into a number (NC) of separate classes EC_i, $(i = 1, ..., NC)$. For each such a class the mean value of the considered error is computed as well as the number of cases of the SIFR (i.e. computed within the given class EC_i sum of the values of $sifr$ as recorded in Step 9). Fig 1 presents exemplary results obtained for problems where $n = 4, k = 4$, number of classes is $NC = 35$.

The plot (a) in Fig. 1 illustrates the relationship between the averages of aggregated absolute error AAE and the probabilities of SIFR, while the one labeled (b) shows the relationship between aggregated relative error ARE and the probabilities od SIFR. The most important fact is that both the relationships are monotonic - the greater the magnitude of an error (AE or RE) the greater the chances for the SIFR. Although it looks intuitive, it is not an obvious fact - our studies show that similarly strong relationship cannot be observed when instead of the whole AHP setup as here, one focuses on a single prioritization problem (i.e. separately for one fixed criterion)

What is also quite surprising, in difference to the suggestions in literature (see [[12]), it seems that the absolute errors AE manifest stronger monotonic relationship with the probabilities of SIFR, than the relative ones. This shows that the heuristic reasoning underlying some conclusions may be sometimes misleading. The relationship illustrated in Fig 1 was obtained for the case where the number of alternatives is 5 and the number of criteria is 3. However, the same, in spirit, relationship is observed for all other examined numbers of

TABLE I: Values of Spearman correlation coefficients between the probabilities (frequencies) of SIFR and the mean values of AAE (first number) and ARE (second number). Only fractional digits are presented.

$n \setminus k$	3	4	5	6	7
3	978/964	978/941	980/954	942/901	919/948
4	988/961	962/920	955/924	960/952	970/941
5	955/918	970/886	982/940	986/938	969/958
6	990/925	976/904	983/880	993/901	983/879
7	987/941	982/893	985/910	983/902	992/950
8	976/878	974/896	995/861	986/876	995/967

TABLE II: Values of Pearson correlation coefficients between the probabilities (frequencies) of SIFR and the mean values of AAE (first number) and ARE (second number). Only fractional digits are presented.

$n \setminus k$	3	4	5	6	7
3	973/965	980/906	972/941	946/894	927/937
4	974/949	957/903	951/901	934/912	959/858
5	956/900	952/865	942/907	958/926	952/958
6	973/915	976/863	953/880	975/954	975/868
7	970/938	974/908	976/910	975/870	985/942
8	932/865	959/892	983/835	978/867	973/947

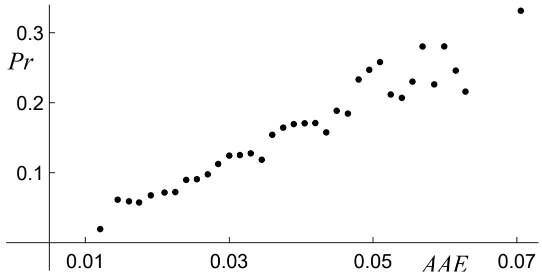
alternatives and criteria. In our studies we have considered $n = 3, ..., 8$ and $k = 3, ..., 7$. The degree of monotonicity of the considered relationships can be measured with the help of Spearman correlation coefficients. Table I shows their values obtained in all considered cases. Column heads indicate the number of criteria k, heads of rows indicate the number of alternatives. To save the table space, in appropriate cells only fractional digits are presented.

We can see that apparently both types of errors are strongly correlated with the probability of SIFR, and in that sense both are meaningful. However in all cases the Spearman correlation coefficient computed for the AAE is greater than the one computed in case of the ARE. So the results confirm our previous conclusion. Moreover, one can also notice that the correlation of ARE decreases when the number of alternatives increases, while the correlation of AAE is much more robust against such changes.

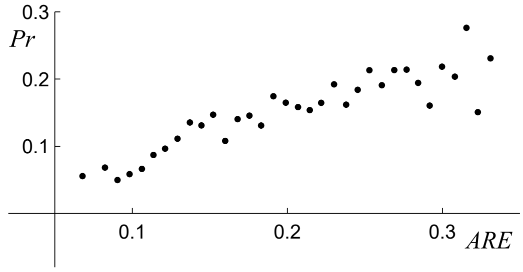
Apart from the monotonicity, it is also interesting to learn whether or not the relationships are similar to the linear one (as sugested by the 1. It appears to be true for all investigated cases, the values of corresponding Pearson correlation coefficients are presented in Table II). Again, the linearity of the relationships is a bit stronger in the case of the absolute errors AE.

Next issue that is important for the PCM acceptance procedure is the notion of small errors. As it was indicated in the previous section our proposal is to consider as "small" PVEEs such ones that are of similar magnitude as those resulting solely from the rounding procedure. So, our another task is to determine how small are the "small" errors and what is their relationship with the adopted judgment scales. For this study we assume the following simulation framework:

Step 0 (Initialization) Set: n - the number of alternatives, k- the number of criteria, N -the number of simulated AHP problems, the prioritization method (GM or REV)
Step 1 Randomly generate the true priority vectors $\mathbf{v}^i,\ i =$



(a)



(b)

Fig. 1: The relationship between the probabilities (frequencies) of SIFR (labeled as Pr) and the AAE, plot (a), and ARE, plot (b). This exemplary graph was obtained in the case where $n = 5,\ k = 3$. The whole range of observed errors was split into $NC = 35$ classes.

$1, ..., k$, and compute related perfect comparison matrix $\mathbf{M}^i$ with elements $m^i_{j,l} = \frac{v^i_j}{v^i_l}$

Step 2 For every considered judgment scale separately, compute rounded matrices $\mathbf{RM}^i, i = 0, ..., k$, by rounding all values in the upper triangle of $\mathbf{M}^i, i = 0, ..., k$, to the closest value from the scale and replace all elements in the lower triangle of the $\mathbf{RM^i}$ with the reciprocities of the appropriate elements from the upper triangle.

Step 3 With the help of adopted prioritization method compute the values of the estimates of the vectors $\mathbf{v}^i,\ i = 1, ..., k$ along with the errors A_i, R_i, AAE or ARE. Write down values computed in this step as one record.

Step 4 N times repeat Steps 1 and 3

Step 5 Return all records organized as one database.

In our studies we compare the Saaty's scale (SS), Extended Saaty's scale ESS[17], and geometric scale GS[2], (their definitions are provided in Section 2). Results of our simulation studies are summarized in Table III

The usage of a judgment scale and thus the rounding procedure is an immanent step of all pairwise comparison judgments. Consequently, the rounding errors cannot be avoided and have to be accepted. So it is natural to treat each error in judgment that has similar magnitude to the rounding error as a small one or maybe even as one that is negligible. In our opinion the limit for this kind of errors should be set by the mean of the observed rounding errors. If we accept this point, then we can notice that in the light of the results presented in Table III the magnitude of this limit for small errors depends on the assumed judgment scale. And from this point of view both the GS[2] and EES[17] looks much better than the usual SS. Such a poor performance of the SS were also pointed out in other research conclusions, see e.g. [26].

Another important observation is that even such small errors may lead to SIFR! It is another interesting and not obvious fact revealed by presented here studies. If we, for example, take into account the relationship illustrated in Fig 1, we can see that AA errors of magnitude less that the limits for small errors are related to AHP problems where the probability of SIFR is above 0.05 (see the results for the SS in Table III) . The coefficients presented in Table II were computed in simulations where the GM prioritization method was used to obtain the EPV. However, when we use the REV method the results are basically the same and they lead to the same conclusions, so we omit their presentation..

TABLE III: Selected statistics for AA errors that are results of the rounding procedure - dependence on the adopted judgment scale. The considered scales are SS, ESS[17] and GS[2] as defined in Section 2.

Statistics		Min	Max	Mean	St. Deviat.
Scale:	n				
SS	3	0.0053	0.0357	0.0154	0.0041
SS	4	0.0050	0.0325	0.0136	0.0036
SS	5	0.0048	0.0300	0.0123	0.0033
SS	6	0.0043	0.0276	0.0111	0.0030
SS	7	0.0041	0.0254	0.0101	0.0027
SS	8	0.0041	0.0233	0.0092	0.0025
ESS[17]	3	0.0042	0.0304	0.0140	0.0037
ESS[17]	4	0.0038	0.0237	0.0114	0.0027
ESS[17]	5	0.0039	0.0218	0.0098	0.0023
ESS[17]	6	0.0035	0.0196	0.0086	0.0020
ESS[17]	7	0.0033	0.0170	0.0077	0.0018
ESS[17]	8	0.0031	0.0162	0.0070	0.0017
GS[2]	3	0.0033	0.0208	0.0106	0.0025
GS[2]	4	0.0034	0.0211	0.0084	0.0021
GS[2]	5	0.0025	0.0206	0.0073	0.0020
GS[2]	6	0.0025	0.0173	0.0065	0.0019
GS[2]	7	0.0022	0.0174	0.0059	0.0017
GS[2]	8	0.0019	0.0145	0.0053	0.0016

V. Final remarks

The simulation experiments described in this paper revealed several interesting facts. First fact is that not the relative errors but the absolute errors are better correlated with the chances of significantly incorrect final ranking. Second, that the small estimation errors - i.e. the ones of magnitude similar to the rounding errors - are not negligible because they also may cause the change of order of the two most important alternatives, and the probability of such situation is between 5% and 8%. Next interesting observation is that adoption of the geometric scale (here the GS[2]) leads to smaller rounding errors than the extended Satty's scale (here ESS[17]). The worst one with respect to this criterion is the most commonly used Saaty's scale.

All the issues considered here where investigated with the help of simulation frameworks that take into account the whole hierarchical structure that occurs in the AHP. The conventional simulation approach is to investigate separately a single prioritization problem (i.e. problem of estimating weights on the basis of a fixed and only one PCM), see e.g.

as in [11], [13], [6], [20], [4], [28]. The approach adopted here to simulation analysis of AHP problems was introduced for the first time in [17]. However in the context of error analysis, this approach was proposed for the first time in our paper [14]. What is very important in this context, the relationship between the estimation errors and the chances of SIFR are much more unclear and weak when we consider the single prioritization problems separately. It is an important observation. During our simulation experiments we have observed that it is possible that there are no wrong rankings obtained for every separate prioritization problem within given hierarchy, and the only incorrect ranking is the final one!

So, the presented here results are important because they enrich our knowledge about the prioritization problems in general, revealing the nature of crucial relationships. But they also should have impact on the PCM acceptance methodology. The impact should be at least twofold. First, we know when the PCM should be (or even has to be) accepted; when the AE errors (or RE errors) are small (as defined here). Secondly, the very close relationship between AE errors and the probability of significantly wrong final PV (as illustrated e.g. by Fig. 1 or Table I) can form a sound fundament for development of a new PCM-acceptance procedure that would really on the analysis of the chances of good/ bad consequences of the decisions, and be well justified by the mathematical-statistics methodology.

Finally, let us note that the PVEEs (of both types) cannot be observed directly, but the values of inconsistency indices can be observed instead. The relationship between the values of inconsistency indices and the AAE and ARE errors where analyzed in [12]. It appears that the best correlation was found between the ATI index and the PVEEs of both types. More detailed discussion on this issue can be found in [12]. In view of presented here results further thorough studies within this area should be conducted in future. But in the light of presented results it is very promising direction.

REFERENCES

[1] Aguaron, J., Moreno-Jimenez J.M., (2003). "The geometric consistency index: Approximated thresholds." Euro. J. Oper. Res. 147, 137-145.

[2] Alonso, J., Lamata, T., (2006). "Consistency in the Analytic Hierarchy Process: a New Approach." Inter. J.Uncertain., Fuzzin. Knowl.-Based Syst., 14, 445-459.

[3] Aydin, S., Kahraman, C. (2012). "Evaluation of E-commerce Website Quality Using Fuzzy Multi-criteria Decision Making Approach". IAENG International Journal of Computer Science. 39. 64-70.

[4] Budescu, D.V., Zwick R., Rapoport A. (1986) "Comparison of the analytic hierarchy process and the geometric mean procedure for ratio scaling", Appl. Psychol. Meas., 10, 69-78.

[5] Chamodrakas, I., Batis, D., Martakos, D. (2010). "Supplier selection in electronic marketplaces using satisficing and fuzzy AHP". Expert Systems With Applications,37, 490-498

[6] Choo, E.U., Wedley, W.C. (2004) "A common framework for deriving preference values from pairwise comparison matrices", Comp. Oper. Res. 31, 893-908.

[7] Crawford, G., Williams, C.A. (1985) "A note on the analysis of subjective judgment matrices". J. Math. Psychol. 29, 387-405.

[8] Dong Y., Xu, Y., Li, H., Dai, M.(2008) "A comparative study of the numerical scales and the prioritization methods in AHP", Euro. J. Oper. Res. 186, 229-242.

[9] Farkas A. (2011), "Multi-Criteria Comparison of Bridge Designs", Acta Polytech. Hung., Vol. 8, No.1

[10] Franek, J., Kresta, A. (2014). "Judgment scales and consistency measure in AHP". Procedia Economics and Finance, 12, 164-173. https://doi.org/10.1016/S2212-5671(14)00332-3

[11] Grzybowski, A. Z., "Note on a new optimization based approach for estimating priority weights and related consistency index", Expert Systems with Applications (2012), 39, 11699-11708

[12] Grzybowski, A.Z., "New results on inconsistency indices and their relationship with the quality of priority vector estimation", Expert Systems With Applications 43 (2016) 197-212, ISSN: 0957-4174,

[13] Grzybowski A. Z., Starczewski T.: "Remarks about Inconsistency Analysis in the Pairwise Comparison Technique", Proceedings of IEEE 14th International Scientific Conference on Informatics, Poprad, (eds. Novitzka, S. Korecko, A. Szakal), New York 2017, pp.227-231, ISBN 978-1-5386-0888-3

[14] Grzybowski A. Z., Starczewski T.: "Simulation analysis of prioritization errors in the AHP and their relationship with an adopted judgement scale", Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering and Computer Science 2018, 23-25 October, 2018, San Francisco, USA, pp 47-51,

[15] Kahraman, C., Kaya, I. (2010). "A fuzzy multicriteria methodology for selection among energy alternatives ", Expert Systems With Applications,, Vol. 37, No 9, 6270-6281

[16] Kazibudzki, P.T., "Redefinition of triad's inconsistency and its impact on the consistency measurement of pairwise comparison matrix", Journal of Applied Mathematics and Computational Mechanics, vol. 15, no. 1, pp. 71-78, 2016

[17] Kazibudzki P.T. Grzybowski, A.Z., "On some advancements within certain multicriteria decision making support methodology", American Journal of Business and Management, Vol. 2, No. 2, 2013, 143-154 DOI: 10.11634/216796061302287 (ISSN 2167-9606 (Print) ISSN 2167-9614 (Online)/ Copyright World Scholars, USA)

[18] Koczkodaj W.W. (1993) "A new definition of consistency of pairwise comparisons." Mathematical and Computer Modelling 18(7),79-84.

[19] 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, DOI: 10.3846/20294913.2016.1210694

[20] Lin, C-C. (2007) "A Revised Framework for Deriving Preference Values from Pairwise Comparison Matrices". Euro. J. Oper. Res., 176, 1145-1150.

[21] Saaty, T.L. "Scaling method for priorities in hierarchical structures", (1977) J. Math. Psychol., Vol. 15, No. 3, 234-281.

[22] Saaty, T.L. (1980) The Analytic Hierarchy Process, McGraw Hill, New York.

[23] Saaty, T.L., Vargas, L.G. (1984) "Comparison of eigenvalue, logarithmic least square and least square methods in estimating ratio". J. Math. Model., 5, 309-324.

[24] Saaty, T.L., Vargas L.G.(2013) The Logic of Priorities: Applications of the AHP in Business, Energy, Health and Transportation, RWS Publications, Pittsburgh (ebook)

[25] Starczewski, T. (2016). "Relationship between priority ratios disturbances and priority estimation errors". Journal of Applied Mathematics and Computational Mechanics, 15(3), 143-154. doi:10.17512/jamcm.2016.3.14

[26] Starczewski, T. (2017). "Remarks on the impact of the adopted scale on the priority estimation quality". Journal of Applied Mathematics and Computational Mechanics, 16(3), 105-116. doi:10.17512/jamcm.2017.3.10

[27] Supriya M, Sangeeta K, and G K Patra, "Trustworthy Cloud Service Provider Selection using Multi Criteria Decision Making Methods," Engineering Letters, vol. 24, no.1, pp1-10, 2016

[28] Zahedi, F. (1986) "A simulation study of estimation methods in the analytic hierarchy process", Socio-Econ. Plann. Sci. 20, 347-354.

[29] Zyoud S. H., Fuchs-Hanusch D. (2017). "A bibliometric-based survey on AHP and TOPSIS techniques", Expert Systems With Applications, 78, 158-181.