



HAL
open science

An Empirical Examination of the Consistency Ratio in the Analytic Hierarchy Process (AHP)

Valery Lukinskiy, Vladislav Lukinskiy, Boris Sokolov, Darya Bazhina

► **To cite this version:**

Valery Lukinskiy, Vladislav Lukinskiy, Boris Sokolov, Darya Bazhina. An Empirical Examination of the Consistency Ratio in the Analytic Hierarchy Process (AHP). IFIP International Conference on Advances in Production Management Systems (APMS), Sep 2021, Nantes, France. pp.477-485, 10.1007/978-3-030-85914-5_51 . hal-03897864

HAL Id: hal-03897864

<https://inria.hal.science/hal-03897864>

Submitted on 14 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

An Empirical Examination of the Consistency Ratio in the Analytic Hierarchy Process (AHP)

Valery Lukinskiy^{1,*}, Vladislav Lukinskiy¹, Boris Sokolov², Darya Bazhina¹

¹ National Research University Higher School of Economics (HSE University), St. Petersburg, Russia

² Saint Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, St. Petersburg, Russia
lukinskiy@mail.ru

Abstract. Papers related to solving practical problems in all areas of economic activity use the AHP methodology from the 1980s, proposed on the basis of studying a limited number of tangible objects. One possible option to overcome this situation could be a new approach based on probability theory and mathematical statistics. We used information obtained from processing 292 matrices collected between 2014 and 2020 and reflecting expert judgements. We were able to compare a particular expert's judgement with an array of expert judgements from previous studies to refine the computational operations of the AHP. It was found that the probability estimates of the consistency indices for expert judgements differ significantly from the probability estimates for the generated values; e.g., for 6×6 matrices, they differ by a factor of 20. The results of our study do not confirm the need to revise expert judgements if the consistency ratio (C.R.) is greater than 0.10, and confirm the need for more example calculations for tangible objects and for new dependencies to estimate intangible objects based on artificial intelligence models and methods.

Keywords: Expert Judgements, Consistency Ratio, Tangible Objects

1 Introduction

Ivanov and Sokolov [1] emphasize that during the supply chain (SC) planning horizon, different structural elements (decision-makers, processes, products, control variables, constraints, goals, perturbations, and so on) are involved in the decision making for SC planning. Dolgui et al. [2] highlight that different problems need different solutions, and no one technique can be universally applied to the same effect. According to Felice [3], one of the most commonly used multi-criteria decision making methods is the Analytic Hierarchy Process (AHP). Csató [4] notes that in real-world applications the judgements of decision-makers may be inconsistent: for example, alternative A is two times better than alternative B, alternative B is three times better than alternative C, but alternative A is not six times better than alternative C. Sarraf and McGuire [5] note that the consistency ratio (C.R.) is the ratio of the consistency

index to the average random index for the same order matrix. A C.R. of 0.10 or less is considered acceptable.

In works related to practical tasks (the multi-objective facility layout problem, selection of automated guided vehicles for handling materials, supplier assessment, selection of the most suitable managers for projects, design operations in distribution centres, calculating a supply chain performance index, inventory classification, making sustainable investment decisions, and many others), the value $C.R. \leq 0.1$ is used to assess matrix consistency [6-20]. Table 1 provides information from some of these sources in more detail.

Table 1. Assessing matrix consistency in works related to solving practical problems.

Reference	Main purpose of the study	Remarks
[8]	The paper proposes a model for supplier assessment	As $CR < 0.10$, the degree of consistency is satisfactory. If the value is higher, that would indicate that the PCM must be revised
[11]	This paper proposes a tool for the decision-making process of selecting the most suitable managers for projects	If CR is higher than 0.10, then the judgment matrix is inconsistent and judgments should be reviewed and improved
[17]	AHP and VIKOR techniques are used to evaluate the supplier	If $CR \leq 0.1$, the judgement matrix is acceptable; otherwise, it is considered inconsistent
[20]	The aim of this paper is to provide a framework for sustainable SC development	The value of CR must be lower than 0.10 for an acceptable level of consistency

A number of papers question the need to revise expert opinions if $C.R. > 0.1$. Falsini et al. [21] outline that in the original version of AHP, indeed, if C.R. is greater than 0.10 the decision maker traditionally should not tolerate the error and should reject the analysed matrix. In a business environment, problems connected to this limitation are clearly identifiable: every iteration of the same step implies a cost in terms of time and money. Ishizaka and Labib [22] note that for all consistency checking, some questions remain: what is the cut-off rule to declare my matrix inconsistent? Should this rule depend on the size of the matrix? The results of Ishizaka and Siraj [23] experimentally invalidate the threshold of $C.R. < 0.1$ and suggest a much higher threshold of acceptance. Out of 50 AHP participants, 5 did not report their level of inconsistency in the given questionnaire. Out of the remaining 45 participants, only 8 participants passed the widely accepted criterion of $C.R. < 0.1$. Chao et al. [24] indicate that inconsistency of PCM exists naturally in many sport activities, thus the ranking should depend on the existing matches in spite of the inconsistencies. Abastante et al. [25] declare that the C.R. threshold has to be suggested by the analyst depending on the complexity of the problem. Brunelli [26] emphasizes that not all the facets of pairwise comparison matrices can be fruitfully analyzed by means of randomly generated matrices. This suggests increasing the use of empirical preferences in the study of preference relations. According to Grzybowski and Starczewski [27], expert systems are designed to assist decision-makers and they are not substitutes for them.

Expert systems do not have such human capabilities as feelings, responsibility or simply concern about risk. Thus the final decision about the rejection or acceptance of the PCMs should belong to the DMs.

Thus, works related to practical problem solving in all areas of economic activity use $C.R. \leq 0.1$ to assess matrix consistency [28, 29] in spite of the results of studies on the development of AHP. It should be noted that the value $C.R. \leq 0.1$ was derived from a study on a limited number of tangible objects. One possible way to overcome this situation is to take a new approach based on probability theory and mathematical statistics.

This paper presents a probabilistic approach for evaluating matrix consistency, which, compared to the deterministic approach, allows establishing quantitative relationships between expert and generated matrices. We used information obtained from processing 292 matrices collected between 2014 and 2020 and reflecting expert judgements. The obtained results do not confirm the need to revise expert judgements if $C.R.$ is greater than 0.10, calls for many more example calculations for tangible objects and a search for new dependencies to estimate intangible objects based on artificial intelligence models and methods.

The rest of the paper is organised as follows. The results of designing a methodology for assessing matrix consistency are presented in Section 2. Section 3 provides the directions for future research.

2 Conducting a Probabilistic Assessment of Matrix Consistency

Despite considerable progress in the study and use of the analytic hierarchy process (AHP), some aspects remain under-researched. One of them is the problem of assessing the difference between generated and expert (expert-generated) matrices for tangible and intangible objects.

The choice of a particular criterion value z , equal to a certain value of the consistency ratio $C.R.$, is of fundamental importance, as it affects the rest of the calculation procedure according to the AHP. According to Saaty [28, 29], if the calculated value of $C.R.$ is less than or equal to z , the matrix is considered consistent; in this case, the priority vector w_i and the consistency index $C.I.$ can be used for decision making. If $C.R. > z$, then experts must reconsider some of the judgements in the matrix.

Table 2 shows the statistical characteristics used to determine the parameters of the distribution laws chosen to approximate the $C.I.$ samples of expert and generated matrices. The samples of the first type are a summary of the material presented in various sources (International Journal of Production Research, International Journal of Production Economics, Expert Systems with Applications, European Journal of Operational Research, Decision Support Systems, Omega, and others) and in our own research. We used information obtained from processing 292 matrices collected between 2014 and 2020 and reflecting expert judgements. Samples of the second type were generated on the basis of generated matrices of dimensions $n \times n$ [28, 29].

Table 2. Statistical characteristics and distribution law parameters for expert and generated C.I. samples.

Matrix	Type of sample	Sample size	Statistical characteristics		Distribution law	Parameters of distribution	
			\bar{x}_i	σ		m	b_m
3 × 3	1	76	0.050	0.037	Weibull	1.35	0.73
	2	100	0.382	0.517	Weibull	0.73	0.53
4 × 4	1	60	0.160	0.140	Weibull	1.15	0.95
	2	600	0.894	0.620	Weibull	1.45	0.91
5 × 5	1	86	0.445	0.240	Rayleigh	2.00	0.89
	2	600	1.124	0.540	Truncated normal		
6 × 6	1	57	0.561	0.405	Weibull	1.40	0.91
	2	600	1.225	0.450	Normal		
7 × 7	1	13	0.315	0.251	Weibull	1.25	0.93
	2	600	1.354	0.343	Normal		

Analysis of the statistical data in Table 2 shows that most of the distribution functions of the first and second type of samples can be approximated by the Weibull distribution [30]

$$F(x) = 1 - \exp\left[-\left(\frac{b_m \cdot x}{x}\right)^m\right], \quad (1)$$

where m and b_m are, respectively, shape and scale parameters, and x is an argument (consistency index C.I.).

The parameters are calculated using a method of moments. One can also use tables [30] where the values m and b_m are listed depending on a coefficient of variation v . For instance, for an expert sample of the first type (3 × 3 matrix) at $v = 0.74$, we will get $m = 1.35$, $b_m = 0.73$. Then the probability of consistency at $z = 0.05$ and, consequently, C.I. = 0.026 will amount to

$$F(C.I. = 0.026) = 1 - \exp\left[-\left(\frac{0.73 \cdot 0.026}{0.050}\right)^{1.35}\right] = 0.237$$

Table 3 shows the results of the $F(C.I.)$ distribution functions calculated for the first and second type samples.

Table 3. Probabilistic assessment of expert F_1 (C.I.) and generated F_2 (C.I.) matrices.

Matrix	C.I.	C.R. $\leq z$	F_1 (C.I.)	F_2 (C.I.)	F_1 (C.I.)/ F_2 (C.I.)
3×3	0.026*	0.050*	0.237	0.1500	1.58
	0.052	0.100	0.600	0.2300	2.61
	0.100	0.193	0.900	0.350	2.57
4×4	0.050	0.056	0.281	0.0140	20.07
	0.072*	0.080*	0.313	0.0230	13.61
	0.100	0.112	0.423	0.0350	12.09
5×5	0.110*	0.100*	0.070	0.0270	2.59
	0.200	0.180	0.140	0.0400	3.50
6×6	0.125*	0.100*	0.101	0.0052	19.42
	0.200	0.160	0.187	0.0080	23.37
7×7	0.135*	0.100*	0.270	0.0002	1350
	0.200	0.149	0.420	0.0004	1050

*Extreme values of C.I. for this sample

An analysis of Table 3 (excluding 7×7 matrices due to the small sample size of the first type) shows that

- probability estimates F_1 (C.I.) corresponding to the limit value of C.I. ranged from 0.07 (5×5 matrix) to 0.313 (4×4 matrix), i.e., more than 31% of the expert matrices are consistent;
- for the generated matrices, the F_2 (C.I.) values ranged from 0.15 (3×3 matrix) to 0.0052 (6×6 matrix), i.e., almost 30 times lower;
- F_1 (C.I.)/ F_2 (C.I.) ratio varies widely from 1.58 (3×3 matrix) to 19.42 (6×6 matrix).

A possible reason for these contradictions is that the average C.I. of the generated matrices is chosen as the random consistency index R.I., which is generated based on a uniform distribution without observing the transitivity requirements between the row and column judgements. Therefore, making the correct decision regarding the limits of C.R. ≤ 0.1 values based on the generated matrices requires additional verification of the tangible matrices.

Let's consider a possible approach to solving this problem by comparing the areas of five geometric figures: a circle, triangle, square, diamond, and rectangle [31]. The priority vectors calculated based on expert opinion (w_i) and actual data on the areas of geometric figures (w_{i_ST}) are shown in Table 4 [31].

We consulted three experts, the results of whose calculations are also shown in Table 4. In addition to these indicators, we calculated the average values of absolute deviations according to the formula

$$\Delta = \frac{\sum_{i=1}^n |w_i - w_{i_ST}|}{n}, \quad (2)$$

where w_i is the value of the i -th element of the priority vector obtained by the experts (priorities from comparisons); w_{i_ST} is the value of the i -th element of the priority vector obtained from actual data on the areas of geometric figures (actual relative size).

Table 4. Results of calculations of priority vectors w_i and average values of absolute deviations Δ .

Figures	w_{i_ST}	Priority vector, w_i			
		Saaty [31], C.R.=0.005	Expert 1, C.R.=0.036	Expert 2, C.R.=0.036	Expert 3, C.R.=0.179
Circle	0.471	0.457	0.485	0.555	0.539
Triangle	0.050	0.049	0.052	0.041	0.030
Square	0.234	0.257	0.233	0.185	0.227
Diamond	0.149	0.150	0.152	0.140	0.143
Rectangle	0.096	0.087	0.078	0.079	0.061
Average deviation, Δ		0.0096	0.0076	0.034	0.027

The following conclusions can be drawn from the analysis of Table 4:

- when C.R. values are equal, the average values of absolute deviations of evaluations of the first and second experts differ by a factor of almost 4.5 ($\Delta = 0.0076$ and $\Delta = 0.034$);
- for the third expert (C.R. = 0.179), the average value of absolute deviation is $\Delta = 0.027$, which is less than the average value of the absolute deviation for the second expert ($\Delta = 0.034$), for whom C.R. = 0.036;
- the value of the absolute deviation of the estimates made by the expert [31] is $\Delta = 0.0096$, which is higher than the value for the first expert ($\Delta = 0.0076$).

To illustrate the proposed approach, let us consider a case where a manufacturing company is predicted to purchase 1,000 units of five component parts, the number of which is proportional to the actual data on the areas of geometric figures w_{i_ST} (Table 4). Given the uncertainty of demand, four experts, whose estimates are proportional to the priority vectors w_i (Table 4), were engaged to clarify the volume of purchases of each of the nomenclatures.

The results of the calculation of the discrepancies of the expert and predicted values are given in Table 5. For example, Table 5 shows that if the procurement was based on the judgement of Expert 1, the total number of shortage and surplus units would be 38 units. Taking into account the €10 penalty for each deficit or surplus unit, the additional cost to the company would be €380. The corresponding calculations for the other experts are shown in Table 5, from which it follows that the forecast variant based on the priority vector of Expert 1 (C.R. = 0.036) is the preferable option.

The paradoxical nature of these results calls for many more example calculations for tangible objects and a search for new dependencies to estimate intangible objects based on artificial intelligence models and methods.

Table 5. Evaluating the results of a product purchase decision.

Component parts	w_{i_ST}	Discrepancies between predicted and expert estimates			
		Saaty [31], C.R.=0.005	Expert 1, C.R.=0.036	Expert 2, C.R.=0.036	Expert 3, C.R.=0.179
Number 1	471	14	14	84	68
Number 2	50	1	2	9	20
Number 3	234	23	1	49	7
Number 4	149	1	3	9	6
Number 5	96	9	18	17	35
Discrepancies, in units		48	38	168	136
Discrepancies, in euros		480	380	1680	1360

3 Conclusion

A probabilistic approach for evaluating matrix consistency has been proposed, which, compared to the deterministic approach, allows establishing quantitative relationships between expert and generated matrices. Analysis of the results from processing 292 matrices shows that the probabilistic estimates of consistency indices (in the case of expert judgements) differ significantly from the probabilistic estimates of consistency indices for the generated values. For example, for 6×6 matrices, the estimates differ by a factor of 20. The results of our study do not confirm the need to revise expert judgements if C.R. is greater than 0.10.

The main directions for further research are the following:

1. The matrix generation algorithm should be refined with the introduction of artificial intelligence elements, which can be identified from the analysis of tangible object matrices.
2. Collection, systematisation, and statistical analysis of expert judgements should be continued in order to obtain representative samples of matrices. They could then be compared with artificial (generated) matrices, which would provide reliable multi-criteria estimates.

Acknowledgments

The research described in this paper is partially supported by state research 0073–2019–0004.

References

1. Ivanov, D., Sokolov, B.: Structure dynamics control approach to supply chain planning and adaptation. *International Journal of Production Research* 50(21), 6133-6149 (2012)

2. Dolgui, A., Ivanov, D., Sokolov, B.: Ripple effect in the supply chain: an analysis and recent literature. *International Journal of Production Research* 56(1-2), 414-430 (2018)
3. Felice, F.: Research and applications of AHP/ANP and MCDA for decision making in manufacturing. *International Journal of Production Research* 50(17), 4735-4737 (2012)
4. Csató, L.: A characterization of the Logarithmic Least Squares Method. *European Journal of Operational Research* 276, 212–216 (2019)
5. Sarraf, R., McGuire, M.: Integration and comparison of multi-criteria decision making methods in safe route planner. *Expert Systems With Applications* 154, 113399 (2020)
6. Singh, S., Singh, V.: Three-level AHP-based heuristic approach for a multi-objective facility layout problem. *International Journal of Production Research* 49(4), 1105-1125 (2011)
7. Maniya, K., Bhatt, M.: A multi-attribute selection of automated guided vehicle using the AHP/M-GRA technique. *International Journal of Production Research* 49(20), 6107-6124 (2011)
8. Dai, J., Blackhurst, J.: A four-phase AHP-QFD approach for supplier assessment: a sustainability perspective. *International Journal of Production Research* 50(19), 5474-5490 (2012)
9. Salgado, E., Salomon, V., Mello, C.: Analytic hierarchy prioritisation of new product development activities for electronics manufacturing. *International Journal of Production Research* 50(17), 4860-4866 (2012)
10. Azizi, M., Azizpour, M.: A BOCR structure for privatisation effective criteria of Iran newsprint paper industry. *International Journal of Production Research* 50(17), 4867-4876 (2012)
11. Varajão, J., Cruz-Cunha, M.: Using AHP and the IPMA Competence Baseline in the project managers selection process. *International Journal of Production Research* 51(11), 3342-3354 (2013)
12. Quezada, L., López-Ospina, H.: A method for designing a strategy map using AHP and linear programming. *Int. J. Production Economics* 158, 244–255 (2014)
13. Luthra, S., Mangla, S., Xu, L., Diabat, A.: Using AHP to evaluate barriers in adopting sustainable consumption and production initiatives in a supply chain. *Int. J. Production Economics* 181, 342–349 (2016)
14. Vieira, J., Toso, M., Silva, J., Ribeiro, P.: An AHP-based framework for logistics operations in distribution centres. *International Journal of Production Economics* 187, 246–259 (2017)
15. D’Inverno, G., Carosi, L., Romano, G., Guerrini, A.: Water pollution in wastewater treatment plants: An efficiency analysis with undesirable output. *European Journal of Operational Research* 269, 24–34 (2018)
16. Katiyar, R., Meena, P., Barua, M., Tibrewala, R., Kumar, G.: Impact of sustainability and manufacturing practices on supply chain performance: Findings from an emerging economy. *International Journal of Production Economics* 197, 303–316 (2018)
17. Awasthi, A., Govindan, A., Gold, S.: Multi-tier sustainable global supplier selection using a fuzzy AHP-VIKOR based approach. *International Journal of Production Economics* 195, 106–117 (2018)
18. Ishizaka, A., Lolli, F., Balugani, E., Cavallieri, R., Gamberini, R.: DEASort: Assigning items with data envelopment analysis in ABC classes. *International Journal of Production Economics* 199, 7–15 (2018)
19. Vega, D., Vieira, J., Toso, E., Faria, R.: A decision on the truckload and less-than-truckload problem: An approach based on MCDA. *International Journal of Production Economics* 195, 132–145 (2018)

20. Mastrocinque, E., Ramírez, F., Honrubia-Escribano, A., Pham, D.: An AHP-based multi-criteria model for sustainable supply chain development in the renewable energy sector. *Expert Systems With Applications* 150, 113321 (2020)
21. Falsini, D., Fondi, F., Schiraldi, M.: A logistics provider evaluation and selection methodology based on AHP, DEA and linear programming integration. *International Journal of Production Research* 50 (17), 4822–4829 (2012)
22. Ishizaka, A., Labib, A.: Review of the main developments in the analytic hierarchy process. *Expert Systems with Applications* 38, 14336–14345 (2011)
23. Ishizaka, A., Siraj, S.: Are multi-criteria decision-making tools useful? An experimental comparative study of three methods. *European Journal of Operational Research* 264, 462–471 (2018)
24. Chao, X., Kou, G., Li, T., Pen, Y.: Jie Ke versus AlphaGo: A ranking approach using decision making method for large-scale data with incomplete information. *European Journal of Operational Research* 265, 239–247 (2018)
25. Abastante, F., Corrente, S., Greco S., Ishizaka, A., Lami, I.: A new parsimonious AHP methodology: Assigning priorities to many objects by comparing pairwise few reference objects. *Expert systems with applications* 127, 109–120 (2019)
26. Brunelli, M.: A study on the anonymity of pairwise comparisons in group decision making. *European Journal of Operational Research* 279, 502–510 (2019)
27. Grzybowski, A., Starczewski, T.: New look at the inconsistency analysis in the pairwise-comparisons-based prioritization problems. *Expert Systems with Applications* 159, 113549 (2020)
28. Saaty, T.: *The Analytic Hierarchy Process*. McGraw-Hill, New York (1980)
29. Saaty, T.: *Fundamentals of Decision Making and Priority Theory with the Analytical Hierarchy Process*. RWS Publications, Pittsburgh, PA, USA (1994)
30. Weibull, W.: *Fatigue testing and analysis of results*. Pergamon press, Oxford, London (1961)
31. Saaty, T.: Neurons the decision makers, Part I: The firing function of a single neuron. *Neural Networks* 86, 102–114 (2017)