TJOR#1824553, VOL 0, ISS 0

Introduction and comparative analysis of the multi-level parsimonious AHP methodology in a public transport development decision problem

Szabolcs Duleba

QUERY SHEET

This page lists questions we have about your paper. The numbers displayed at left are hyperlinked to the location of the query in your paper.

The title and author names are listed on this sheet as they will be published, both on your paper and on the Table of Contents. Please review and ensure the information is correct and advise us if any changes need to be made. In addition, please review your paper as a whole for typographical and essential corrections.

Your PDF proof has been enabled so that you can comment on the proof directly using Adobe Acrobat. For further information on marking corrections using Acrobat, please visit http://journalauthors.tandf.co.uk/production/acrobat.asp; https://authorservices.taylorandfrancis.com/how-to-correct-proofs-with-adobe/

The CrossRef database (www.crossref.org/) has been used to validate the references.

AUTHOR QUERIES

Q1 The citation of Table 2 has been changed as Table 3. Please correct if this is inaccurate.

Q2 Tables 11–15 were not cited in the text so a citation has been inserted. Please provide a correction if this is inaccurate.

Q3 A disclosure statement reporting no conflict of interest has been inserted. Please correct if this is inaccurate.

Q4 References "Awasthi et al. (2018), Bagherinejad and Fakhrzad (2019), Barradale and Cornet (2018), Bozóki and Fülöp (2018), Bozóki and Tsyganok (2017), Chen et al. (2017), Ignaccolo et al. (2017), Kamaruzzaman et al. (2018), Pedroso et al. (2018), Saaty (1990), Shiau et al. (2015), Sun et al. (2019), Xue et al. (2019), and Zhang et al. (2019)" are listed in the references list but are not cited in the text. Please either cite the references or remove them from the references list.

Q5 Please note that the ORCID section has been created from information supplied with your manuscript submission/CATS. Please correct if this is inaccurate.

JOURNAL OF THE OPERATIONAL RESEARCH SOCIETY
https://doi.org/10.1080/01605682.2020.1824553

ORIGINAL ARTICLE

Introduction and comparative analysis of the multi-level parsimonious AHP methodology in a public transport development decision problem

Szabolcs Duleba

Budapest University of Technology and Economics, Budapest, Hungary

ABSTRACT
The methodology of Parsimonious Analytic Hierarchy Process (PAHP) has been originally constructed to unburden the evaluators of an AHP survey from the numerous pairwise comparisons caused by the several alternatives in decision problems. So far, there are very few applications of PAHP and all of them referred to the last level of the decision structure; to the alternatives. This paper aims to demonstrate a multi-level PAHP based model, so how to apply the method on an arbitrary level of the decision structure. Since being a new method, another objective is to conduct a comparative analysis on the correlation of the results between AHP and multi-level PAHP models. Moreover, the new multi-level approach makes it possible to demonstrate an immanent analysis originated from AHP logic to test PAHP results, which is unique in the scientific literature for Parsimonious technique. The created model has been tested in the real-world decision problem of a Turkish big city, Mersin. Based on the test and analysis, the method can be a suitable tool in case of layman evaluations even in case of small number of criteria or alternatives. In the paper, a general application process is also proposed for other future appliers of the method.

ARTICLE HISTORY
Received 16 December 2019
Accepted 12 September 2020

KEYWORDS
Parsimonious Analytic Hierarchy Process (PAHP); comparative analysis; public transport; multi criteria decision making

1. Introduction

In every preference type survey, the researchers face the dilemma of either acquiring the most valuable data as possible by the questionnaire, asking many and complex questions, or keeping the evaluations as simple as possible for reaching high response rate. The Analytic Hierarchy Process (AHP) is an example for the first option, the questionnaire is complex, containing several pairwise comparisons, while the evaluation time is long and severe cognitive effort is required from the respondents. Simple Additive Weighting (SAW) or Multi-attribute Value Theory (MAVT) represent the second option avoiding pairwise comparisons and hierarchy in their process.

However, the AHP has a clear advantage compared to any other survey methods. It provides a hierarchical decision structure to the respondents, thus leads them by the questions and also motivates the consistent evaluation by the checking procedure. It is an asset in case of layman respondents because the risk of getting confusing or uncertain scoring is reduced. This paper presents a decision problem in which to keep the nature of AHP has been a clear objective because of the public participants. Further, even due to the non-expert evaluators, the application of a more simple and understandable questionnaire has also been among the objectives. Completing both of these – at first sight – contradictory goals has led the author of this paper to create the multi-level Parsimonious AHP model.

Parsimonious AHP (PAHP) is a recently created (Abastante et al., 2018) methodology which aims to unburden decision makers in the Analytic Hierarchy Process by requiring less pairwise comparisons than in the classical AHP procedure. As being a fresh technique, there are only few applications of it and some questions still remained open in terms of the conditions and limitations related to the usage of PAHP.

This paper aims to investigate the question; how to apply PAHP for a multi-level decision problem in which public respondents participate in the survey. Multi-level decision problems are very common in AHP applications, this is one of the advantages of the method that it can handle complex problems with several levels of criteria. All the PAHP models so far investigated the method on alternatives (which are obviously on the same, last level of the decision structure) the recent objective is to examine the applicability on criteria in an arbitrary level of the decision tree.

There are only two articles currently available in the scientific literature introducing and applying the PAHP technique. The first related paper (Abastante et al., 2018) published in this topic, demonstrated the method for evaluating social housing project initiatives in which a board (Programma Housing) - connected to the Italian Bank Foundation - played the role of the

CONTACT Szabolcs Duleba duleba.szabolcs@mail.bme.hu Department of Transport Technology and Economics, Faculty of Transportation Engineering and Vehicle Engineering, Budapest University of Technology and Economics, Műegyetem rkp. 3, Budapest, Hungary.

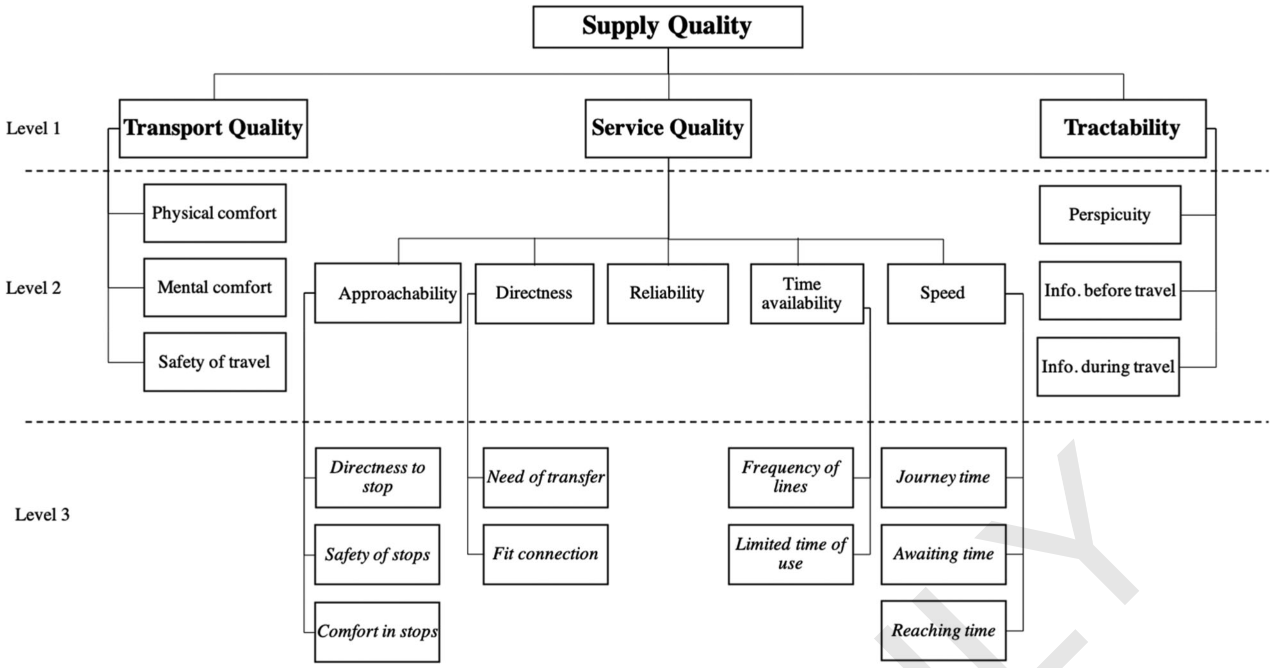

Figure 1. The hierarchical decision structure for public transport development (*Source*: Duleba et al., 2012).

decision maker. Although Programma Housing as an operating entity consisted experts from several different fields (e.g. engineers, architects, financial experts, psychologists), the board was considered as one homogeneous decision maker group and the overall result was gained by making no distinction among evaluators. In this demonstrative application, 10 criteria was selected (five social and five technical) in a uni-level decision structure (not including the level of alternatives). The other known example is the paper of Abastante et al. (2019), in which the effectiveness of PAHP was examined by an experiment with 100 university students. During the frames of this scientific project, students were asked to estimate the area of different geometrical figures and the evaluation was made both by classical and parsimonious AHP. In this study, the authors also provided an example in which a dean selected appropriate students based on their estimated results of three subjects by PAHP. The weights of the subjects can be considered as criteria level but there were no sub-levels except for the alternatives.

This paper endeavors to fill this scientific gap by examining the application of PAHP for a multi-level decision problem. Based on the survey experience, this application is not trivial and a slightly modified model has to be created for gaining the appropriate results. Moreover, a previous complete AHP study from 2018 has been available, so the direct comparison of AHP and multi-level PAHP results has also been possible. Finally, an unprecedented immanent comparison is also demonstrated for getting deeper insight to the PAHP methodology and outcome.

Although parsimonious Analytic Hierarchy Process was originally created for decision problems with many alternatives (Abastante et al., 2018), the core objective of the method also exists in case of many decision criteria. In the demonstrated case study of public transport system development, 24 hierarchical criteria describe the supply quality of public bus service. That means that for a classical entire AHP procedure $(23 \times 24)/2 = 276$ pairwise comparison evaluations would be necessary if all criteria would be on the same level. Considering the created three levels and seven branches of the examined decision problem (see Figure 1), the number of the required pairwise comparisons for AHP is 27. Note that elements of the decision tree are not compared if they are situated on different levels or branches. For public participants completing even this number of comparisons (27) might be too demanding if they are not committed enough to the actual development. The largest comparison matrix is 5×5 in the decision structure and the tolerably consistent filling of this sized matrix requires significant cognitive effort from non-expert evaluators. Thus, urban transport system improvement is a suitable case study for demonstrating multi-level PAHP.

In the followings, first the scientific background of the examined public transport decision problem is overviewed highlighting primarily the public participation. In further parts of the manuscript are introduced AHP, PAHP and modified PAHP (the proposed method). Also, the results of the AHP and the modified PAHP surveys conducted in a Turkish city, Mersin, are demonstrated. Then two types of comparative analysis are presented: an immanent check, which is unprecedented in the Parsimonious methodology, and a direct rank correlation test between AHP and PAHP results. In the Conclusions section the characteristics of PAHP and modified PAHP are analyzed and compared to classical AHP applications. Moreover, a general multi-level PAHP procedure is suggested for the arbitrary application of this new multi-criteria decision making technique.

2. Literature review on multi-level criteria AHP and public involvement in transport planning

Evidently, the multi-level criteria decision making models in MCDM and within that in AHP are dominant. Zhang et al. (2018) applied a two-level (main criteria and sub-criteria) model to prioritize factors of geothermal energy utilization technologies. Ameen and Mourshed (2019) used four criteria levels for a sustainability framework in an AHP model. Kaur et al. (2018) even applied a five criteria level decision hierarchy for authorship verification and compromised account detection in online social networks. Corrente et al. (2012) stated directly, that in theory, the great majority of multi-criteria decision adding (MCDA) methods consider evaluation criteria at the same level, however, practical applications generally impose a hierarchical structure of criteria.

However, when a multi-level criteria decision system is applied, it obviously requires more cognitive efforts from the decision makers (Saaty, 1977). If the participants possess deep insight into the examined problem, have the enthusiasm to participate and are selected experts of the relevant field, it is expected from them to tackle this complexity of the decision. But how can this be expected from common people representing the public opinion without any professional background in the field? Thus, a different approach has to be applied for these problems. An option to overcome the complexity issue might be to use linguistic terms instead of the general scaling of AHP, however, in this case, the number of comparisons remain (Luo et al., 2019).

There are specific cases in which stakeholder involvement is recommended to reach a sustainable decision. Generally, implementations concerning public services should take into consideration ex-ante the preferences of the public. These problems are considered as group decision-making and require the aggregation of preferences by a possibly fair aggregation technique (Aouni et al., 2018; Lin et al., 2018; Wu et al., 2018). Although the demonstrative case study, the urban transport system development is a relevant example of group decision making, the recent paper aims to introduce the Parsimonious AHP technique for hierarchical criteria, thus the emphasis is on the decision hierarchy.

Le Pira et al. (2017) collected six types of methodologies that can be applied for involving public stakeholders to a public transport development decision. Three of them were based on consultation and participation i.e. Focus groups, Multi-actor MCDM methods (Macharis & Bernardini, 2015), PPGIS (Jankowski, 2009), while the others on stakeholder analysis. Parsimonious AHP has also a consultation element in its decision procedure (Abastante et al., 2019) and with citizens as respondents this characteristics is proven to be very useful.

In his AHP survey, Lupo (2013) found that non-expert respondents, in that case customers, provided uncertain judgements on transport services. He proposed a combination of fuzzy theory and analytic hierarchy process in order to mitigate the uncertainty. For multi-level evaluations of public services by citizens this is a common problem (Bai et al., 2008). Except for using fuzzy sets, another possible method for reducing uncertainty of public respondents within the AHP framework is the application of Interval AHP models (Ghorbanzadeh et al., 2018). In some cases, the fuzzy approach is combined with a hesitant linguistic methodology both for criteria and alternative evaluation (Peng & Xia, 2019).

The risk connected to common people participants in AHP surveys refers to two basic issues: incomplete pairwise comparisons and inconsistent pairwise comparisons. An evaluation is incomplete when missing values can be detected in the filled questionnaires independently from being purposive or not. Some methodological solutions exist for this problem (Bozóki et al., 2010; Chen et al., 2015; Duleba et al., 2012) but all of them provide only estimation on the missing elements of the matrices by some optimization procedures.

There are some remarkable techniques to measure and handle inconsistency in AHP such as Kulakowski (2018) or Maturo et al. (2019). Despite, Saaty's Consistency Ratio is still widely applied and accepted (Farooq & Moslem, 2019) and the most recommended technique to handle the problem is the direct modification of the non-tolerably inconsistent matrix by the dialogue with the original evaluator.

Methodologically, Parsimonious AHP and Sparse AHP have got common roots. Sparse AHP refers to decision problems in which only a limited amount of information is available (Oliva et al., 2017) and not all the values of pair wise comparisons. For these cases, Sparse AHP applies the Sparse Eigenvector Method (SEM) (Oliva et al., 2019) in which the unknown utilities are approximated via the dominant eigenvector even if perturbations occur. There are other papers in AHP literature coping with missing entries of the ratio matrix e.g. Fedrizzi and Giove (2007), Menci et al. (2018). The

Table 1. The structure of a $(q \times q)$ consistent, theoretical PCM.

w^1/w^1	w^1/w^2	…	…	w^1/w^q
w^2/w^1	w^2/w^2	w^2/w^3	…	…
…	…	…	…	…
…	…	…	…	…
w^q/w^1	w^q/w^2	…	….	w^q/w^q

main difference between Sparse and Parsimonious AHP is that the reference points in PAHP are systematically selected and further interactions between the respondents and survey instructors are provided compared to Sparse AHP.

Parallel to Sparse AHP, the already mentioned incomplete AHP shows some similarity to Parsimonious AHP from methodological point of view. This is also a pairwise comparison reduction, since only some of the required elements of the PCM-s are available for the analyzers (Benítez et al., 2015; Subramanian & Ramanathan, 2012). The missing elements might be due to mistakes or the lack of confidence of the evaluators or even due to impossibility as in Bozóki et al. (2016) in which tennis players were compared by their head to head results and some of them could not have such results because of their different age. Abastante et al. (2019) stressed that based on Carmone et al. (1997) in incomplete PCM-s the maximum reduction of the pairwise comparisons is by 50% of the total evaluations. Moreover, the incomplete cases do not reflect the question that which elements should be provided in order to gain a feasible final decision on the examined problem.

Parsimonious AHP can be a solution to all the above listed gaps in AHP surveys with public respondents. It reduces the cognitive effort requirement in a multi-level decision structure, by its consultative characteristics (it is included in the steps of PAHP) it mitigates the risk of incomplete evaluations to zero and in case of non-tolerably consistent filling of pairwise comparison matrices, by a dialogue with the respondents' values can be modified to reach the 0.1 threshold value of the Consistency Ratio.

In the followings, the methodological basis of AHP and PAHP is introduced and a model for multi-level PAHP is proposed.

3. Methodology

3.1. Overview on analytic hierarchy process

Since the created model applies AHP for the non-Parsimonious levels of the decision hierarchy and the Parsimonious calculation also contains an AHP step by checking the direct evaluations, first the AHP is described.

Analytic Hierarchy Process is heavily based on the decision structure created from the decision criteria of the complex decision problem (Saaty, 1977). Criteria, sub-criteria, sub-sub-criteria, etc. are identified with the last level of the alternatives in a decision tree. The linkages of the elements are also important since they determine not only the pairwise comparisons in the procedure, but also the final weight and alternative scores by the consideration of the respective previous level element score.

The method of AHP can be summarized as follows. Let us assume that there are a alternatives in the decision: $A_1\ A_2, \ldots, A_a$, and m criteria: $C_1,\ C_2, \ldots,\ C_m$. Let us denote $\mathbf{A}^{(i)}$ the pairwise comparison matrix of the alternatives with respect to the criterion i, and $v^{(i)}$ the eigenvector calculated from the matrix $\mathbf{A}^{(i)}$ by Saaty's eigenvector method (other possible methods also exist). Let also have $\mathbf{C}$, the pairwise comparison matrix of the criteria and $w^{(c)}$ is the weight vector gained from the computed eigenvector of $\mathbf{C}$. Then the final evaluation scores (denoted by $u(w)$) of the alternatives can be gained by:

$$u(w) = v_1^{(C)} w^1 + v_2^{(C)} w^2 + \ldots + v_m^{(C)} w^m \qquad (1)$$

The eigenvector method for PCM-s applies the following calculation for $\mathbf{C}$ and $\mathbf{A}^{(i)}$:

$$\mathbf{A}\ w = \ \lambda_{\max}\ v^i. \qquad (2)$$

Then eigenvector v can be calculated as

$$(\mathbf{A}\ -\ \lambda_{\max}\ .\mathrm{I})\ v^i = 0, \qquad (3)$$

where λ_{max} is the maximum eigenvalue of the matrix $\mathbf{A}$.

Note that the consistent, theoretical pairwise comparison matrices complete the transitivity criterion of the matrix elements: $a_{ik} = \ a_{ij}a_{jk}$ for any i,j,k denoting the rows and columns of the matrix. It is evident, that the principal eigenvector of the consistent matrix exhibited in Table 1 is $\{v^1;v^2;v^3;\ v^4;\ldots;\ v^q\}$ with the maximum eigenvalue of q.

In AHP the eigenvector is calculated just as the PCM-s were consistent. Theoretically they are, but for the experiential matrices (filled by real evaluators) the eigenvectors are probably not consistent. Since the eigenvector calculation is exactly the same for the experiential (most likely not consistent) matrices, an inconsistency threshold is applied in Analytic Hierarchy Process for filtering out (Brunelli & Fedrizzi, 2019) those evaluations which are over the tolerable inconsistency value (see Formula 4 and 5 below). In AHP this value is 0.1 for the Consistency Ratio (defined by Formula 5).

In the AHP procedure, the respondents fill the provided PCM-s by providing their estimation on the relative importance/significance and scoring the brackets of the PCM above the main diagonal (due to

Table 2. Pairwise scale (1-9) of relative weight of criteria in PCMs (Saaty, 1994).

Numerical values	Verbal scale of importance
1	Equal importance of both elements
2	Equally to moderately important
3	Moderate importance of one element over another
4	Moderately to strongly important
5	Strong importance of one element over another
6	Strongly to very strongly important
7	Very strongly important
8	Very strongly to extremely important
9	Extreme importance of one element over another

the reciprocity scoring other positions is redundant). They select values from the Saaty scale (Table 2).

For the experiential PCM-s, the reciprocity ($a_{ji} = 1/a_{ij}$, where $a_{ii} = 1$) is necessary to be fulfilled, however, these experiential matrices are most likely not consistent in terms of the perfect transitivity of the pairwise evaluations, so: $a_{ik} \neq a_{ij}a_{jk}$, where *i,j* and *k* denote arbitrary rows and columns of the PCM.

Consequently the consistency of evaluators' rating have to be measured in experiential matrices of **C** and $\mathbf{A}^{(i)}$. For this, at first Random Index (RI) has to be determined by selecting the value of the same size PCM as investigated (Table 3).

Then Consistency Index (CI) is calculated by using Equation (4).

$$\mathrm{CI} = (\lambda_{max} - n)/(n-1) \qquad (4)$$

where λ_{max} is the principal eigenvalue of the PCM and n is the number of rows in the comparison matrix.

Having gained the RI and CI values, the Consistency Ratio (CR) can be calculated by

$$\mathrm{CR} = \mathrm{CI}/\mathrm{RI}. \qquad (5)$$

CR is acceptable when its value is smaller than 0.1 (Aczel & Saaty, 1983; Saaty, 1994).

There are many approaches for aggregating evaluators' answers, the most popular aggregation procedure is the geometric mean (Aczel & Saaty, 1983).

If "*h*" evaluators take part in the procedure

$$f(x_1, x_2, \ldots, x_h) = \sqrt[h]{\prod_{d=1}^{h} x_d}. \qquad (6)$$

where $x_1, x_2, \ldots, x_h$ denotes entries, in the same position of pairwise comparison matrices, filled in by the decision makers ($d = 1, \ldots, h$).

For determining the final weight vectors of the criteria, the following method is generally applied:

$$w_{A_i} = \frac{w_j}{w}\,\frac{v_{ij}}{\sum_{k=1}^{n} v_{ik}} = \left(\frac{w_j}{w}\,\frac{1}{\sum_{k=1}^{n} v_{ik}}\right) v_{ij}, \qquad (7)$$

where $j = 1, \ldots, m$ and $w = \sum_{i=1}^{m} w_j$;
$wj > 0$ ($j = 1, \ldots, m$) represents the related weight coordinate from the previous level; $v_{ij} > 0$ ($i = 1, \ldots, n$) is the eigenvector computed from the matrix in the current level, w_{Ai} ($i = 1, \ldots, n$) is the calculated weight score of current level's elements. Sensitivity analysis enables in understanding the effects of changes in the main criteria on the sub criteria ranking and help the decision maker to check the robustness throughout the process.

Table 3. Random Index (RI) indices from randomly generated matrices.

N	1	2	3	4	5	6	7	8
RI	0	0	0.58	0.9	1.12	1.24	1.32	1.41

Conducting the sensitivity analysis is also part of the AHP procedure which enables the decision makers to check the robustness of the results by detecting the impact of slight changes of certain weight scores on the whole decision structure ranking.

3.2. The parsimonious AHP approach

The general PAHP approach for C_m criteria and *a* alternatives created by Abastante et al. (2018) can be summarized in the following steps:

1. Direct rating of the criteria C_m with respect to alternative *a* on a given scale (e.g. 0-100);
2. Selecting some reference criteria. Let us denote the reference criteria by *s*, and the number of reference criteria by t_m;
3. Following Corrente et al. (2016), evaluators are therefore asked to apply the original AHP approach to the set composed of the reference evaluations (let us denote γ_{ms} the direct evaluation of the *s* reference element by the evaluators) defined on Step 2 obtaining the normalized evaluations $\mathrm{u}(\gamma_{ms})$, for all $m = 1, \ldots, n$ and for all $s = 1, \ldots, t_m$;
4. The following items have to be checked and discussed with the evaluators:

- the consistency of the pairwise comparisons has to computed by the consistency ratio (CR),
- the normalized evaluations $\mathrm{u}(\gamma_{ms})$, for all $m = 1, \ldots, n$ and for all $s = 1, \ldots, t_m$, have to be compared with the corresponding ratings $rm(\gamma ms)$ controlling that the monotonicity is satisfied, that is verifying that $\mathrm{rm}(\gamma_{ms1}) > \mathrm{rm}(\gamma_{ms2})$ if $\mathrm{u}(\gamma_{ms1}) \geq \mathrm{u}(\gamma_{ms2})$,
- evaluators must have the possibility to modify the rating or the PC-s in order to get consistent pairwise comparisons that satisfy monotonicity which would be accepted by the evaluators.
- The ranking of the evaluations provided by evaluators (which are not the reference evaluations) are obtained by linear interpolation for the normalized evaluations which have to be computed in the previous step. Let us denote:

$u(r_m\ (a))$ the normalized evaluation score of criterion (m) with respect to alternative (a), which is computed from interpolation of the values obtained by the reference AHP in the third and fourth steps,

$u(\gamma_{ms})$ the normalized AHP score of the s reference element gained from pairwise comparison,

$u\ (\gamma_{ms+1})$ the normalized AHP score of the $s+1$ reference element gained from pairwise comparison,

γ_{ms} the direct evaluation of the s reference element by the evaluators (level wisely direct rating),

γ_{ms+1} the direct evaluation of the $s+1$ reference element by the evaluators (level wisely direct rating),

$r_m(a)$ the direct corresponding rating provided by the evaluator to the criteria (a), (level wisely direct rating).

For each $rm(a)\ \in\ [\gamma_{ms},\ \gamma_{ms+1}]$, the following value has to be computed (by this, the rank reversal problem is totally avoided):

$$u(r_m\ (a)) =\ u(\gamma_{ms}) + \frac{u\ (\gamma_{ms+1}) - u\ (\gamma_{ms})}{\gamma_{ms+1}\ -\ \gamma_{ms}}\ (r_m(a) -\ \gamma_{ms}). \tag{8}$$

Basically, in Parsimonious AHP we have a monotonic increasing order of the criteria, determined by direct evaluation on a certain level. From this order we select t_m reference elements and conduct the pairwise comparisons on them. Their increasing order has to be kept and consistency to be checked. Afterwards, we substitute their weight values with the direct ratings and get back to the original monotonic increasing order. In this order normalizing any r_j elements means to find its low (γ_{ms}) and high (γ_{ms+1}) neighbor out of the reference elements. Their monotonic order cannot be changed, which is assured by formula (8).

In the followings, the parsimonious model for multi-level criteria is introduced.

3.3. The proposed multi-level parsimonious AHP technique

Let us have m criteria structured in a decision problem into l levels. Thus we have $k = 1, \ldots,\ l$ levels in the decision, $k\epsilon K$, and the m criteria is distributed to the l levels. Note that in this model there are no alternatives applied, only the weights of the criteria are important to be determined. Let us denote p the criteria on a certain level of the decision, so c_{kp} denotes a criterion on a certain level in which if we have g criteria, $p = 1, \ldots,\ g$. $p\epsilon M$ and M is the set of all criteria of the decision, thus $M = 1, \ldots,\ m$. Consequently, c_{kp} denotes the p-th criterion of the k-th level, so that, for example c_{11} is the first criterion of the first level in the hierarchical structure of the decision.

The first step of the suggested method is to select level or levels in the decision for which the Parsimonious AHP will be conducted. It is proposed to select that level(s) which have $g \geq 9$ that can be considered as enough number worth unburdening the evaluators from numerous pairwise comparisons. Moreover, it is recommended (following Saaty's 7 ± 2 rule for a PCM) to select level(s) for which larger or equal to 5×5 pairwise comparison matrices should be evaluated. Based on own experience the pair wise comparisons for a 5×5 matrix might be demanding for layman evaluators.

As second step, direct evaluations have to be made for the chosen level(s) for the c_{kp} criteria with respect to the goal of the decision problem on a scale from 0 to 100.

Then s reference elements have to be selected on the chosen level(s) k.

Afterwards, the original AHP pairwise comparisons are conducted for the chosen criteria, obtaining the normalized AHP scores for the s criteria: $u(c_{ks})$, for all $s = 1, \ldots, t_p$.

Following the PAHP procedure by Abastante et al. (2018), consistency and monotonicity are being checked and the required modifications are made.

Finally, the formula (8) is applied for all criteria (c_{kp}) existing on the Parsimonious level(s) using:

$$u(c_{kp}\ ) =\ u(c_{ke}) + \frac{u\ (c_{ke+1}) - u\ (c_{ke})}{\gamma_{ke+1}\ -\ \gamma_{ke}}\ (r_{kp} -\ \gamma_{ke}). \tag{9}$$

With respect to $e = 1, \ldots, t_p$ and c_{kp} has the importance between the two reference criteria c_{ke} and c_{ke+1}, thus $r_{ke} < r_{kp} < r_{ke+1}$. Having finished with the Parsimonious level(s) the decision structure should be reconstructed in order to gain the final weight and alternative scores and ranking. Consequently, all $u(c_{kp})$-s have to be multiplied (see Formula 7) by the weight score of its respective element from the previous level k-1. Also, due to the characteristics of AHP, for the lower levels, the new $u(c_{kp})$ weight scores have to be applied for multiplying the scores of the respective lower elements.

3.4. The immanent analysis of the parsimonious scores

In AHP, due to its hierarchical logic, the total distribution of any upper level decision element score to its connected lower level elements is applied.

Table 4. The attributes of public transport supply quality.

Decision elements (criteria)	Level of decision elements	Description
Service quality	Level 1	All service excluding on-vehicle services and information service
Transport quality	Level 1	Services during the time spent on the public vehicle
Tractability	Level 1	The provided information about the time and spots of the journey
Approachability	Level 2	Line access connected services
Directness	Level 2	Simplicity of reaching the destination without shifting vehicles
Time availability	Level 2	The time frame when using a certain vehicle
Speed	Level 2	The speed of the whole travel process
Reliability	Level 2	On time arrivals, keeping the schedule
Physical comfort	Level 2	The comfort of seats, physical space in the bus, air conditioning
Mental comfort	Level 2	Environmental aspects, the behaviour of the driver and other passengers
Safety of travel	Level 2	The perception of safety, the security of the journey
Perspicuity	Level 2	Comprehension of schedule and information
Information before travel	Level 2	Amount and quality of information previously of the journey
Information during travel	Level 2	Availability and quality of information on the vehicle
Directness to stops	Level 3	Proximity of the starting bus stop
Safety of stops	Level 3	The security of bus stops in terms of road safety
Comfort in stops	Level 3	Roof, heating and cooling systems, seats of the stops
Need of transfer	Level 3	The need to change vehicles to the destination
Fit connection	Level 3	On time connection between bus lines or between other types of public transport (trams, trains, etc.)
Frequency of lines	Level 3	Scheduled and realised frequency of the buses
The limited time of use	Level 3	The time frame between the first and the last line of the day
Journey Time	Level 3	Time spent on the vehicle
Awaiting time	Level 3	Waiting time for the desired line
Time to reach stops	Level 3	Time to reach the departure bus stop

However, if one (or more) levels are evaluated by the Parsimonious approach, this distribution has to be examined and in case of imperfect share, the conclusions has to be drawn.

As first step of the immanent analysis, the final Parsimonious scores (u_{ckp}) have to be normalized on the Parsimonious level.

$$nu_{ckp} = \frac{u_{ckp}}{\sum_{p=1}^{g} u_{ckp}} \tag{10}$$

Then in case of perfect Parsimonious scoring (which is very unlikely) the following condition is fulfilled based on Formula 7.

$$w_j = \sum_{p=1}^{b} nu_{ckp} \tag{11}$$

where w_j is the AHP score of the connected upper level element, $b > 0$ is the number of connected elements to the criterion j. In case b elements on the Parsimonious level connect to the upper level (k-1) criterion p.

All detected differences are considered as sensitivity analysis in the multi-level Parsimonious procedure. In the case study, the results of the immanent analysis are also demonstrated.

4. A case study for testing the multi-level parsimonious AHP model on a public transport development decision

The proposed multi-level Parsimonious AHP model has been tested on a real decision making problem: the possible improvement of a public bus transport system in an emerging city: Mersin, Turkey. From methodological point of view the chosen case study seemed fortunate, since previous experience of this problem has been available, published in Duleba and Moslem (2018, 2019). The applied decision structure of criteria has become the same as used in other different AHP surveys (Duleba, 2019; Duleba et al., 2012, 2013) as exhibited in Figure 1.

The attributes of public transport supply quality can be interpreted by Table 4.

For testing the multi-level Parsimonious AHP model as introduced in the Methodology section, first the Parsimonious level (or levels) has to be created. As visible on Figure 1, the most appropriate level is Level 2., since it contains most of the criteria (11) and there is a 5×5 PCM containing the Approachability, Directness, Reliability, Time Availability and Speed factors. Based on this, the Parsimonious model for decision making can also be constructed (Figure 2).

Figure 2 demonstrates the logic of a multi-level PAHP model. The criteria in the selected Parsimonious level(s) are handled regardless their position in the decision tree, ignoring their branches (exhibited by dotted line), thus their direct evaluation can be conducted. From this phase, all the 5 steps of PAHP can be followed as introduced in the Methodology section (in bold, the reference elements are exhibited for step 2). Having conducted all steps, for computing the final scores even for the Parsimonious level criteria, the structure is rebuilt so branch connections are considered again by multiplying the scores by the respective previous level element scores.

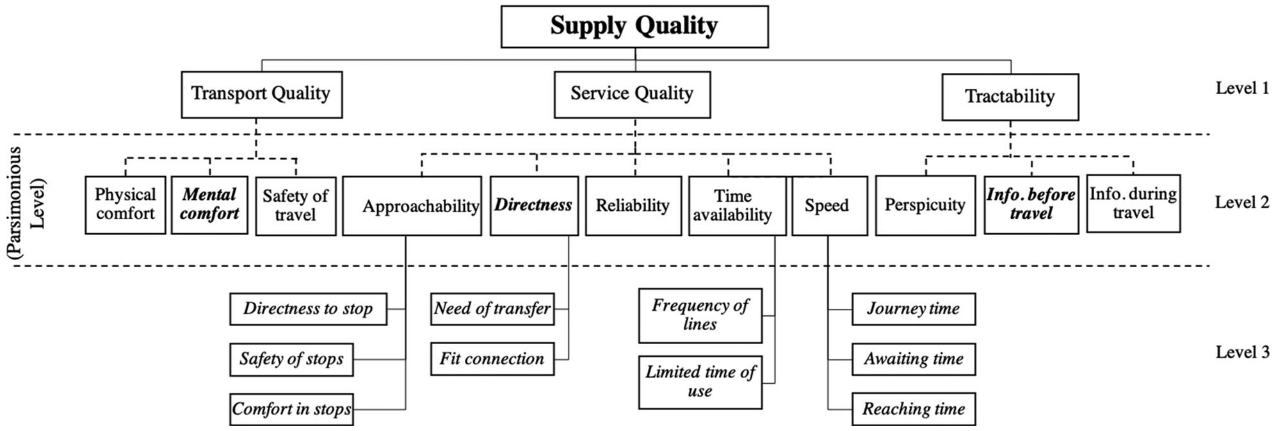

Figure 2. The decision structure reflecting the Parsimonious criteria.

As a real life pattern, 42 passengers were interviewed in the survey in Mersin for the Parsimonious AHP evaluations both the direct and pairwise part. An instructor, Sarbast Moslem PhD student conducted the evaluation process from 29th April to 1st May in 2019. The respondents were selected randomly in the same bus stops as in 2018. Since this procedure contained a consultation phase (see step 4 in the Methodology section) for the case if the pairwise comparisons of the reference criteria were not consistent, the contacts of the evaluators were required for a possible second round interview, in which the PCM values could have been modified for reaching the tolerable inconsistency (CR < 0.1) and desirable monotonicity. Having aggregated and computed the pairwise comparison scoring, it turned out that the second round was not necessary to keep (see step 5 in this section). The final results were discussed with the representatives of the local public bus company and they verified the relevance of the outcomes.

For each interview, approximately 10 min were spent in explaining and filling up the questionnaire and it can be stated that PAHP is far less time consuming than the classical AHP. By using the multi-level PAHP model, altogether 6 PCM-s had to be filled: a 3×3 on the first level, another 3×3 for the Parsimonious level for the reference criteria, and two 3×3 and two 2×2 matrices for the third level elements. All other efforts of the evaluators had just been direct evaluations for the second level criteria which meant a much easier job for them than participating in a classical AHP survey. Taking this decision structure the evaluators should have filled 8 PCM-s including a 5×5 large pairwise comparison matrix. That would have meant altogether 21 pairwise evaluations while in the PAHP survey only 14 had to be done with avoiding the confusing 5×5 PCM. The reflections of the public participants verified the simplicity and the effort and time consuming characteristics of the PAHP process highlighting the better understandability compared to the 2018 AHP survey in which the respondents were complaining about the complexity of the questionnaire.

The following steps have been applied for obtaining the normalized scores in the selected PAHP level (second level);

1. Level 2 has been selected from the decision hierarchy due to the most number of level decision elements (11), moreover avoiding the 5x5 PCM derived from the Service Quality branch.
2. The criteria are directly evaluated (level wisely "level 2") by passenger evaluators based on (0-1) scale.
   The gained normalized increasing order has been: Mental comfort (0.03); Approachability (0.04); Info during travel (0.06); Time availability (0.07); Speed (0.08); Directness (0.09); Safety of Travel (0.10); Perspicuity (0.11); Physical comfort (0.12); Reliability (0.13); Info before travel (0.17).
3. In this phase, appropriate reference criteria have to be selected out of the total criteria in the Parsimonious level:
   For this decision problem, the elements "Information before travel" ($\gamma_1 = 0.17$), "Directness" ($\gamma_2 = 0.09$) and "Mental comfort" ($\gamma_3 = 0.03$) have been chosen based on their positions in the direct evaluations: the highest, a middle and the lowest values out of the order of second level criteria. It has to be emphasized that selecting too many reference criteria would cause the same problem as conventional AHP, a too large pairwise comparison matrix should be evaluated. Obviously in case of several (over 20) criteria on the parsimonious level at least four or five reference points might be necessary, but in this survey, three out of the 11 elements seemed to be sufficient (in the paper of Abastante et al., 2019, the authors chose four reference points out 13 elements) and a 3×3 PCM is not difficult to evaluate even by public respondents. However, the question of the optimal number of reference elements in PAHP

Table 5. The PC of the reference criteria.

CR = 0.09418	Information before travel	Directness	Mental comfort	Scores	Rank
Information before travel	1	3.984	2.994	0.620	1
Directness	0.251	1	1.992	0.224	2
Mental comfort	0.334	0.502	1	0.156	3

Table 6. The direct evaluation of the level 2 criteria by passengers and the final parsimonious scores.

Criteria	Direct evaluation	Rank	Final Parsimonious scores
Approachability	0.04	10	0.167
Directness	0.09	6	0.224
Time availability	0.07	8	0.201
Speed	0.08	7	0.213
Reliability	0.13	2	0.422
Physical comfort	0.12	3	0.373
Mental comfort	0.03	11	0.156
Safety of travel	0.10	5	0.274
Perspicuity	0.11	4	0.323
Info before travel	0.17	1	0.620
Info during travel	0.06	9	0.190

methodology is still open and definitely deserves further research.

4. The reference elements have been compared pair wisely following the conventional AHP approach.
5. Consistency ratio has been calculated and monotonicity has been checked in this phase. Based on the results, another round for interviewing has not been necessary since the CR got a value smaller than 0.1 (CR = 0.09418), which is considered consistent enough and the monotonicity condition has also been completed. (The rank of the three reference criteria remained the same after the pairwise comparisons.)
6. The priorities for all the criteria have been obtained (the comprehensive evaluation was done by employing Formula (7))
   Within the interval all other criteria scores have been repositioned, however the ranking order has been kept for all criteria in the parsimonious level.

$$u(\text{Approachability}) = 0.156 + \frac{0.224 - 0.156}{0.09 - 0.03}(0.04 - 0.03) = 0.156 + \frac{0.068}{0.06}(0.01) = 0.167$$

$$u(\text{Time availability}) = 0.156 + \frac{0.224 - 0.156}{0.09 - 0.03}(0.07 - 0.03) = 0.201$$

$$u(\text{Speed}) = 0.156 + \frac{0.224 - 0.156}{0.09 - 0.03}(0.08 - 0.03) = 0.213$$

$$u(\text{Reliability}) = 0.244 + \frac{0.62 - 0.224}{0.17 - 0.09}(0.13 - 0.09) = 0.244 + \frac{0.396}{0.08}(0.04) = 0.422$$

$$u(\text{Physical comfort}) = 0.244 + \frac{0.62 - 0.224}{0.17 - 0.09}(0.12 - 0.09) = 0.373$$

$$u(\text{Safety of travel}) = 0.244 + \frac{0.62 - 0.224}{0.17 - 0.09}(0.13 - 0.10) = 0.274$$

$$u(\text{Perspicuity}) = 0.244 + \frac{0.62 - 0.224}{0.17 - 0.09}(0.13 - 0.11) = 0.323$$

$$u(\text{Info during travel}) = 0.156 + \frac{0.224 - 0.156}{0.09 - 0.03}(0.06 - 0.03) = 0.19$$

Before determining the final scores in the decision, the original decision structure (see Figure 1) has to be rebuilt. Based on this, the final parsimonious scores have to be normalized by their position in the original hierarchy. For instance, the parsimonious score of 'Approachability', 'Directness', 'Time availability', 'Speed' and 'Reliability' have to be normalized to 1 (Table 7). Moreover the other, non-parsimonious levels and criteria scores have to be attached to the decision (Table 5).
According to the PAHP approach outcomes (Table 6), the final scores and the criteria ranking are presented in Tables 7 and 8.

7. Consequently, prior to gain the final decision scores, the original decision structure (see Figure 1) has to be reconstructed. Logically, the parsimonious scores have to be normalized based on their position in the original hierarchy. Following the decision structure, the parsimonious scores of the elements have to be normalized to one branch wisely (see Table 8). Evidently, the other, non-parsimonious level criteria scores are considered at this stage.
8. Based on this, the final decision scores and ranking can be gained and the multi-level PAHP problem can be solved.

The first level in the decision structure is obviously equals to the AHP scoring.

Table 9 exhibits the final normalized scores for the second level of the decision structure which is calculated as following by multiplying the scores by their branch scores:

Approachability $= 0.136 \times 0.374 = 0.051$
Directness $= 0.183 \times 0.\ 374 = 0.068$
Time availability $= 0.164 \times 0.374 = 0.061$
Speed $= 0.174 \times 0.\ 374 = 0.065$
Reliability $= 0.344 \times 0.422 = 0.129$

Table 7. Level 1 and Level 3 scores obtained from the original AHP approach, Level 2 scores obtained from the PAHP approach.

Level 1		Level 2 (parsimonious level)		Level 3	
Supply Quality		*Service Quality*		*Approachability*	
Service Quality	0.374	Approachability	0.167	Directness to stops	0.464
Transport Quality	0.466	Directness	0.224	Safety of stops	0.293
Tractability	0.16	Time availability	0.201	Comfort in stops	0.243
		Speed	0.213	*Directness*	
		Reliability	0.422	Need of transfer	0.675
		Transport Quality		Fit connection	0.325
		Physical comfort	0.373	*Time availability*	
		Mental comfort	0.156	Frequency of lines	0.702
		Safety of travel	0.274	Limited time of us	0.298
		Tractability		*Speed*	
		Perspicuity	0.323	Journey time	0.265
		Info before travel	0.620	Awaiting time	0.236
		Info during travel	0.190	Reaching time	0.499

Physical comfort $= 0.466 \times 0.465 = 0.216$
Mental comfort $= 0.466 \times 0.194 = 0.091$
Safety of travel $= 0.466 \times 0.341 = 0.159$

Perspicuity $= 0.16 \times 0.285 = 0.046$
Information before travel $= 0.16 \times 0.547 = 0.088$
Information during travel $= 0.16 \times 0.168 = 0.027$

Table 10 exhibits the final normalized scores for the last level of the decision structure which is calculated as following by multiplying the element's scores by their branch scores:

Directness to stops $= 0.136 \times 0.374 \times 0.464 = 0.024$
Safety of stops $= 0.136 \times 0.374 \times 0.293 = 0.015$
Comfort in stops $= 0.136 \times 0.374 \times 0.243 = 0.012$

Need of transfer $= 0.183 \times 0.374 \times 0.675 = 0.046$
Fit connection $= 0.183 \times 0.374 \times 0.325 = 0.022$

Frequency of lines $= 0.164 \times 0.374 \times 0.702 = 0.043$
Limited time of use $= 0.164 \times 0.374 \times 0.298 = 0.018$

Journey time $= 0.174 \times 0.374 \times 0.265 = 0.017$
Awaiting time $= 0.174 \times 0.374 \times 0.236 = 0.015$
Time to reach stops $= 0.174 \times 0.374 \times 0.499 = 0.032$

The sensitivity analysis showed stable ranking, however, slightly modifying the weight score of 'Transport Quality' (−0.04) and 'Tractability' (+0.04), for the second level, a moderate change in the ranking could be detected. The rank of 'Info before travel' improved to 4th and 'Perspicuity' to 9th, while 'Mental comfort' and 'Approachability' lost one position each. For more significant change, a relatively large modification would be necessary in the weight scores of the first and second level. Note that the ranking as a whole has been proved to be non-sensitive despite the slight rank reversal (Table 11).

5. An immanent analysis to compare AHP and multi-level PAHP results

One of the strengths of AHP is the inner logic that the score of each element can be composed by the scores of the respective (branch) elements from the lower level. This is the consequence of the many normalization steps and the applied eigenvector method in the calculation.

The introduced new multi-level PAHP model offers the possibility to check whether the results of the new technique correspond to this immanent logic of AHP. The proportion of the sum of scores of one branch elements has to approximate the weight sore of their respective upper level element. This checking procedure is unprecedented in the scientific literature of PAHP. Obviously, it is impossible to get 100% match because of the different scoring and calculation applied in the Parsimonious level, however, this comparison may serve as a sensitivity analysis or a procedure for the validation of PAHP.

Specifically in this model, the proportional sum of the parsimonious scores of 'Approachability', 'Directness', 'Time availability', 'Speed' and 'Reliability' out of the total parsimonious sum of the second level has to approximate the normalized weight score of their respective upper level element, 'Service Quality'. For 'Transport Quality' and 'Tractability' elements the requirement is the same.

Let us check if the gained parsimonious scores of the second (parsimonious level) decision elements can compose their upper level respective criteria scores (Table 12).

Immanent normalized PAHP score of Service Quality = Sum of Service Quality/Total parsimonious $= 1.227/3.163 = 0.388$

Immanent normalized PAHP score of Transport Quality = Sum of Transport Quality/Total parsimonious= $0.803/3.163 = 0.254$

Immanent normalized PAHP score of Tractability = Sum of Tractability/Total parsimonious $= 1.133/3.163 = 0.358$

Based on the immanent check it can be concluded that the AHP score of 'Service Quality' (0.374) is well approximated by the PAHP score (0.388). However, 'Transport Quality' is strongly underestimated and 'Tractability' is overestimated by the PAHP scoring. This phenomenon can be explained by the dominance of the element 'Info

Table 8. Level 1 and Level 3 scores obtained from the original AHP approach, Level 2 scores obtained from the PAHP approach after normalization.

Level 1		Level 2 (parsimonious level)		Level 3	
Supply Quality		*Service Quality*		*Approachability*	
Service Quality	0.374	Approachability	0,136	Directness to stops	0.464
Transport Quality	0.466	Directness	0,183	Safety of stops	0.293
Tractability	0.16	Time availability	0,164	Comfort in stops	0.243
		Speed	0,174	*Directness*	
		Reliability	0,344	Need of transfer	0.675
		Transport Quality		Fit connection	0.325
		Physical comfort	0,465	*Time availability*	
		Mental comfort	0,194	Frequency of lines	0.702
		Safety of travel	0,341	Limited time of us	0.298
		Tractability		*Speed*	
		Perspicuity	0,285	Journey time	0.265
		Info before travel	0,547	Awaiting time	0.236
		Info during travel	0,168	Reaching time	0.499

Table 9. Final normalized scores by passengers for level 2 (after their normalization by branches).

	Level 2	
Criteria	Scores	Rank
Approachability	0,051	9
Directness	0,068	6
Time availability	0,061	8
Speed	0,065	7
Reliability	0,129	3
Physical comfort	0,216	1
Mental comfort	0,091	4
Safety of travel	0,159	2
Perspicuity	0,046	10
Info before travel	0,088	5
Info during travel	0,027	11

Table 10. Final normalized scores by passengers for level 3 (after their normalization by branches).

	Level 3	
Criteria	Scores	Rank
Directness to stops	0,024	4
Safety of stops	0,0149	9
Comfort in stops	0,012	10
Need of transfer	0,046	1
Fit connection	0,022	5
Frequency of lines	0,043	2
Limited time of use	0,018	6
Journey time	0,017	7
Awaiting time	0,015	8
Time to reach stops	0,032	3

Table 11. Results of the sensitivity analysis after first level weight score modification (rank change marked with bold).

	Level 2	
Criteria	Scores	Rank
Approachability	**0,051**	**10**
Directness	0,068	6
Time availability	0,061	8
Speed	0,065	7
Reliability	0,129	3
Physical comfort	0,198	1
Mental comfort	**0,083**	**5**
Safety of travel	0,145	2
Perspicuity	**0,057**	**9**
Info before travel	**0,109**	**4**
Info during travel	0,036	11

before travel' in the PAHP scoring. As Table 5 demonstrates, in the pairwise comparison of the selected reference criteria, 'Info before travel' received

Table 12. The immanent check of PAHP scores.

Level 1		Level 2 (parsimonious level)	
Supply Quality		*Service Quality*	
Service Quality	0.374	Approachability	0.167
Transport Quality	0.466	Directness	0.224
Tractability	0.16	Time availability	0.201
		Speed	0.213
		Reliability	0.422
		Sum *Service Quality*: 1.227	
		Transport Quality	
		Physical comfort	0.373
		Mental comfort	0.156
		Safety of travel	0.274
		Sum *Transport Quality*: 0.803	
		Tractability	
		Perspicuity	0.323
		Info before travel	0.620
		Info during travel	0.190
After normalization		**Sum *Tractability*:**	**1.133**
***Service quality* 0.388**		**Total parsimonious**	**3.163**
***Transport Quality* 0.254**			
***Tractability* 0.358**			

significantly higher weight score than the other two reference items 'Directness' and 'Mental comfort'. Note that the importance of this element has only been mitigated by the low AHP score of its respective upper level element, 'Tractability'. Even with this unfavorable scoring, it has been ranked 5th in the overall scoring (Table 9). Thus, it can be stated that the significance of this criterion is most likely higher than its final score in the model and it has been revealed by the immanent check of PAHP scores. Note that the sensitivity analysis showed very stable ranking and a 0.04 modification in the first level caused merely one improved position for 'Info before travel' criterion.

In the immanent analysis, 'Tractability' gained score from 'Transport quality' merely because of this single item and this created the difference between the AHP and PAHP scores. Consequently, taking into account the extreme high value of that item, the overall approximation would be much better if threating separately that element.

Table 13. Spearman's Rank correlation coefficient for Level 1.

Criteria	Rank of 2018 AHP survey	Rank of 2019 PAHP survey	d_i	$(d_i)^2$
Service quality	2	2	0	0
Transport Quality	1	1	0	0
Tractability	3	3	0	0
$m=3$		$R=1$		

Table 14. Spearman's Rank correlation coefficient for Level 2.

Criteria	Rank of 2018 AHP survey	Rank of 2019 PAHP survey	d_i	$(d_i)^2$
Approachability	5	9	−4	16
Directness	3	6	−3	9
Time availability	4	8	−4	16
Speed	11	7	4	16
Reliability	8	3	5	25
Physical comfort	2	1	1	1
Mental comfort	7	4	3	9
Safety of travel	1	2	−1	1
Perspicuity	10	10	0	0
Information before travel	6	5	1	1
Information during travel	9	11	−2	4
$m=11$		$R=0.5545$		

Table 15. Spearman's Rank correlation coefficient for Level 3.

Criteria	Rank of 2018 AHP survey	Rank of 2019 PAHP survey	d_i	$(d_i)^2$
Directness to stops	3	4	−1	1
Safety of stops	6	9	−3	9
Comfort in stops	7	10	−3	9
Need of transfer	2	1	1	1
Fit connection	4	5	−1	1
Frequency of lines	1	2	−1	1
Limited time of use	5	6	−1	1
Journey time	9	7	2	4
Awaiting time	10	8	2	4
Time to reach stops	8	3	5	25
$m=10$		$R=0.6606$		

6. A direct comparative analysis

Since the results of a previous AHP survey in the same city using the same decision structure is available, it is advisable to conduct a rank comparison between the two surveys. In both research, 42 evaluators were directly interviewed, however, not the same persons, so different scoring could have be expected even if they evaluated the same public transport system without any changes during the one year difference in time. For the comparison, Spearman's rank correlation technique has been applied. The 'R' value of the calculation shows the degree of correlation among the two different rankings. Above 0.5, the correlation can be considered as strong and positive between the two surveys. The following formula has been applied for the calculation:

$$R = 1 - \left(\frac{6\sum d^2}{m^3 - m}\right) \tag{12}$$

in which d is the difference between the ranks and m is the number of elements to be compared (Tables 13–15).

The direct comparison resulted in strong rank correlation for all three levels. Note that the third level has been affected by the previous levels and even with this impact, over 66% correlation can be detected, which is remarkable. Selecting the first two most important elements is also noticeable in both surveys. The first level ranking is totally identical, and for the second level, both research identified Physical Comfort and Safety of Travel as most significant criteria, for the third, Frequency of lines and Need of Transfer have been seeded as the most important ones.

Thus, it can be concluded that the direct comparative analysis verified the PAHP methodology.

7. Conclusions

Applying Parsimonious AHP model instead of the conventional AHP methodology has caused less evaluation time and cost, better understandability for participants, while resulted a very similar final ranking of the decision criteria, for all levels as the Spearman index indicated (strong rank correlation). The immanent logic of AHP has not been strictly kept, but conducting the checking procedure, a so far hidden dominant criterion could be detected and the outcomes of PAHP could be explained and verified. Thus, it can be suggested that the immanent checking step should be integrated to the Parsimonious AHP methodology in case of multi-level models.

As limitation, it can be stated that the unique immanent analysis revealed that the multi-level PAHP method is highly sensitive to the phase of pairwise comparing the reference criteria of the parsimonious level. Consequently, not only the Consistency Ratio should be checked as suggested so far in the scientific literature of the technique, but also the extreme high or extreme low weight scores gained in this phase should be negotiated with the evaluators to ensure the real intention of scoring.

The extension of the Parsimonious AHP methodology to the multi-level Parsimonious AHP method opens the gate for several other applications since most of the complex decision models are hierarchical. The case study also demonstrated that the parsimonious approach can be integrated to other multi-criteria models (e.g. AHP), for those levels, which require the reduction of pairwise comparisons.

Remarking the further research, many other parsimonious applications are necessary to get familiar with all characteristics of this new methodology. The objective benefits are clear, it provides faster and cheaper survey process and undoubtedly, the survey pattern can more easily be extended by this technique than applying the complex pairwise comparison questionnaire of the conventional AHP. But are the results of PAHP as trustworthy as AHP which has been applied many times by many researchers successfully? This paper merely provided one example but only many other applications can verify the technique ultimately.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research has been supported by the MTA Bolyai Research Scholarship.

ORCID

Szabolcs Duleba http://orcid.org/0000-0002-2367-752X

References

Abastante, F., Corrente, S., Greco, S., Ishizaka, A., & Lami, I. M. (2018). Choice architecture for architecture choices: Evaluating social housing initiatives putting together a parsimonious AHP methodology and the Choquet integral. *Land Use Policy*, *78*, 748–762. https://doi.org/10.1016/j.landusepol.2018.07.037

Abastante, F., Corrente, S., Greco, S., Ishizaka, A., & Lami, I. M. (2019). A new Parsomonious AHP methodology: Assigning priorities to many objects by comparing pairwise few reference objects. *Expert Systems with Applications*, *127*, 109–120. https://doi.org/10.1016/j.eswa.2019.02.036

Aczel, J., & Saaty, T. L. (1983). Procedures for synthesizing ratio judgements. *Journal of Mathematical Psychology*, *27*, 93–102.

Ameen, R. F. M., & Mourshed, M. (2019). Urban sustainability assessment framework development: The ranking and weighting sustainability indicators using analytic hierarchy process. *Sustainable Cities and Society*, *44*, 356–366. https://doi.org/10.1016/j.scs.2018.10.020

Aouni, B., Doumpos, M., Pérez-Gladish, B., & Steuer, R. E. (2018). On the increasing importance of multiple criteria decision aid methods for portfolio selection. *Journal of the Operational Research Society*, *69*(10), 1525–1542. https://doi.org/10.1080/01605682.2018.1475118

~~Awasthi, A., Sayyadi, R., & Khabbazian, A. (2018). A combined approach integrating gap analysis, QFD and AHP for improving logistics service quality. *International Journal of Logistics Systems and Management*, *29*(2), 190–214. https://doi.org/10.1504/IJLSM.2018.089171~~

~~Bagherinejad, J., & Fakhrzad, P. (2019). Selection process of emerging technologies using AHP FMCGP mixed method: Transportation industry of Iran. *International Journal of Applied Management Science*, *11*(1), 55–71. https://doi.org/10.1504/IJAMS.2019.096651~~

Bai, C. H., Lai, F. J., Chen, Y., & Hutchinson, J. (2008). Conceptualizing the perceived service quality of public utility services: A multi-level, multi-dimensional model. *Total Quality Management & Business Excellence*, *19*(10), 1055–1070. https://doi.org/10.1080/14783360802264228

~~Barradale, M. J., & Cornet, Y. (2018). Developing assessment criteria for sustainable transport appraisal. *Transportation Research Record*, *2672*(3), 104–117.~~

Benítez, J., Delgado-Galván, X., Izquierdo, J., & Pérez-García, R. (2015). Consistent completion of incomplete judgements in decision making using. *Journal of Computational and Applied Mathematics*, *290*(15), 412–422. https://doi.org/10.1016/j.cam.2015.05.023

Bozóki, S., Csató, L., & Temesi, J. (2016). An application of incomplete pairwise comparison matrices for ranking top tennis players. *European Journal of Operational Research*, *248*(1), 211–218. https://doi.org/10.1016/j.ejor.2015.06.069

~~Bozóki, S., & Fülöp, J. (2018). Efficient weight vectors from pairwise comparison matrices. *European Journal of Operational Research*, *264*(2), 419–427. https://doi.org/10.1016/j.ejor.2017.06.033~~

Bozóki, S., Fülöp, J., & Rónyai, L. (2010). On optimal completion of incomplete pairwise comparison matrices. *Mathematical and Computer Modelling*, *52*(1–2), 318–333. https://doi.org/10.1016/j.mcm.2010.02.047

~~Bozóki, S., & Tsyganok, V. (2017). The logarithmic least squares optimality of the geometric mean of weight vectors calculated from all spanning trees for (in) complete pairwise comparison matrices. arXiv Preprint arXiv *1701.04265*.~~

Brunelli, M., & Fedrizzi, M. (2019). A general formulation for some in consistency indices of pairwise comparisons. *Annals of Operations Research*, *274*(1–2), 155–169. https://doi.org/10.1007/s10479-018-2936-6

Carmone, F. J., Kara, A., & Zanakis, S. H. (1997). A Monte Carlo investigation of incomplete pairwise comparison matrices in AHP. *European Journal of Operational Research*, *102*(3), 538–553. https://doi.org/10.1016/S0377-2217(96)00250-0

~~Chen, G., Cheung, W., Chu, S. C., & Xu, L. (2017). Transshipment hub selection from a shipper's and freight forwarder's perspective. *Expert Systems with Applications*, *83*, 396–404. https://doi.org/10.1016/j.eswa.2017.04.044~~

Chen, K., Kou, G., Tarn, J., & Song, Y. (2015). Bridging the gap between missing and inconsistent values in eliciting preference from pairwise comparison matrices. *Annals of Operations Research*, *235*(1), 155–175. https://doi.org/10.1007/s10479-015-1997-z

Corrente, S., Greco, S., & Ishizaka, A. (2016). Combining analytic hierarchy process and Choquet integral within non additive robust ordinal regression. *Omega*, *61*, 2–18. https://doi.org/10.1016/j.omega.2015.07.003

Corrente, S., Greco, S., & Słowiński, R. (2012). Multiple criteria hierarchy process in robust ordinal regression. *Decision Support Systems*, *53*(3), 660–674. https://doi.org/10.1016/j.dss.2012.03.004

Duleba, S. (2019). An AHP-ISM approach for considering public preferences in a public transport development decision. *Transport*, *34*(6), 662–671. https://doi.org/10.3846/transport.2019.9080

Duleba, S., Mishina, T., & Shimazaki, Y. (2012). A dynamic analysis on public bus transport's supply quality by using. *Transport*, *27*(3), 268–275. https://doi.org/10.3846/16484142.2012.719838

Duleba, S., & Moslem, S. (2018). Sustainable urban transport development with stakeholder participation, an AHP-Kendall Model: A case study for mersin. *Sustainability*, *10*(10), 3647. https://doi.org/10.3390/su10103647

Duleba, S., & Moslem, S. (2019). Examining Pareto optimality in analytic hierarchy process on real Data: An application in public transport service development. *Expert Systems with Applications*, *116*, 21–30. https://doi.org/10.1016/j.eswa.2018.08.049

Duleba, S., Shimazaki, Y., & Mishina, T. (2013). An analysis on the connections of factors in a public transport system by. *AHP-ISM Transport*, *28*(4), 404–412.

Farooq, D., & Moslem, S. (2019). Evaluation and ranking of driver behavior factors related to road safety by applying analytic network process. *Periodica Polytechnica Transportation Engineering*, *48*(2), 189–195. https://doi.org/10.3311/PPtr.13037

Fedrizzi, M., & Giove, S. (2007). Incomplete pairwise comparison and consistency optimization. *European Journal of Operational Research*, *183*(1), 303–313. https://doi.org/10.1016/j.ejor.2006.09.065

Ghorbanzadeh, O., Moslem, S., Blaschke, T., & Duleba, S. (2018). Sustainable urban transport planning considering different stakeholder groups by an interval-AHP decision support model. *Sustainability*, *11*(1), 9. https://doi.org/10.3390/su11010009

~~Ignaccolo, M., Inturri, G., García-Melon, M., Giuffrida, N., Le Pira, M., & Torrisi, V. (2017). Combining Analytic Hierarchy Process (AHP) with role playing games for stakeholder engagements in complex transport decisions. 20th EURO Working Group on Transportation Meeting, EWGT 2017, Budapest, Hungary.~~

Jankowski, P. (2009). Towards participatory geographic information systems for community-based environmental decision making. *Journal of Environmental Management*, *90*(6), 1966–1971. https://doi.org/10.1016/j.jenvman.2007.08.028

~~Kamaruzzaman, S. N., Lou, E. C. W., Wong, P. F., Wood, R., & Che Ani, A. I. (2018). Developing weighting system for refurbishment building assessment scheme in Malaysia through analytic hierarchy process (AHP) approach. *Energy Policy*, *112*, 280–290. https://doi.org/10.1016/j.enpol.2017.10.023~~

Kaur, R., Singh, S., & Kumar, H. (2018). Authcom: Authorship verification and compromised account detection in online social networks using AHP-TOPSIS embedded profiling based technique. *Expert Systems with Applications*, *113*, 397–414. https://doi.org/10.1016/j.eswa.2018.07.011

Kulakowski, K. (2018). Inconsistency in the ordinal pairwise comparisons method with and without ties. *European Journal of Operational Research*, *270*, 314–327.

Le Pira, M., Marcucci, E., Gatta, V., Ignaccolo, M., Inturri, G., & Pluchino, A. (2017). Towards a decision-support procedure to foster stakeholder involvement and acceptability of urban freight transport policies. *European Transport Research Review*, *9*, 54–68.

Lin, M., Xu, Z., Zhai, Y., & Yao, Z. (2018). Multi-attribute group decision-making under probabilistic uncertain linguistic environment. *Journal of the Operational Research Society*, *69*(2), 157–170. https://doi.org/10.1057/s41274-017-0182-y

Luo, S., Zhang, H., Wang, J., & Li, L. (2019). Group decision-making approach for evaluating the sustainability of constructed wetlands with probabilistic linguistic preference relations. *Journal of the Operational Research Society*, *70*(12), 2039–2055. https://doi.org/10.1080/01605682.2018.1510806

Lupo, T. (2013). Handling stakeholder uncertain judgements in strategic transport service analyses. *Transport Policy*, *29*, 54–63. https://doi.org/10.1016/j.tranpol.2013.04.002

Macharis, C., & Bernardini, A. (2015). Reviewing the use of multi-criteria decision analysis for the evaluation of transport projects: Time for a multi-actor approach. *Transport Policy*, *37*, 177–186. https://doi.org/10.1016/j.tranpol.2014.11.002

Maturo, F., Ventre, V., & Longo, A. (2019). On consistency and incoherence in analytic hierarchy process and intertemporal choices models. In Flaut C., Hoskova-Mayerova, S., Flaut, D. (Eds.), *Models and theories in social systems. Studies in systems, decision and control* (p. 179). Springer.

Menci, M., Oliva, G., Papi, M., Setola, R., & Scala, A. (2018). A suite of distributed methodologies to solve the sparse analytic hierarchy process problem. *2018 European Control Conference*, IEEE.

Oliva, G., Setola, R., & Scala, A. (2017). Sparse and distributed analytic hierarchy process. *Automatica*, *85*, 211–220. https://doi.org/10.1016/j.automatica.2017.07.051

Oliva, G., Setola, R., Scala, A., & Dell'Olmo, P. (2019). Sparse analytic hierarchy process: An experimental analysis. *Soft Computing*, *23*(9), 2887–2898. https://doi.org/10.1007/s00500-018-3401-9

~~Pedroso, G., Bermann, C., & Sanches Pereira, A. (2018). Combining the functional unit concept and the analytic hierarchy process method for performance assessment of public transport options. *Case Studies on Transport Policy*, *6*(4), 722–736. https://doi.org/10.1016/j.cstp.2018.09.002~~

Peng, J.-G., & Xia, G. (2019). A systematic fuzzy multi-criteria group decision-making approach for alternatives evaluation. *Journal of the Operational Research Society*, *70*(9), 1490–1501. https://doi.org/10.1080/01605682.2018.1495995

Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, *15*(3), 234–281. https://doi.org/10.1016/0022-2496(77)90033-5

~~Saaty, T. L. (1990). How to make a decision: The analytic hierarchy process. *European Journal of Operational Research*, *48*(1), 9–26. https://doi.org/10.1016/0377-2217(90)90057-I~~

Saaty, T. L. (1994). How to make a decision: The analytic hierarchy process. *Interfaces*, *24*(6), 19–43. https://doi.org/10.1287/inte.24.6.19

~~Shiau, T. A., Huang, M. W., & Lin, W. Y. (2015). Developing an indicator system for measuring Taiwan's transport sustainability. *International Journal of Sustainable Transportation*, *9*(2), 81–92. https://doi.org/10.1080/15568318.2012.738775~~

Subramanian, N., & Ramanathan, R. (2012). A review of applications of analytic hierarchy process in operations management. *International Journal of Production Economics*, *138*(2), 215–241. https://doi.org/10.1016/j.ijpe.2012.03.036

~~Sun, B., Tang, J., Yu, D., Song, Z., & Wang, P. (2019). Ecosystem health assessment: A PSR analysis combining AHP and FCE methods for Jiaozhou Bay, China. *Ocean & Coastal Management*, *168*, 41–50.~~

Wu, D., Yang, Z., Wang, N., Li, C., & Yang, Y. (2018). An Integrated multi-criteria decision making model and AHP weighting uncertainty analysis for sustainability assessment of coal-fired power units. *Sustainability*, *10*(6), 1700. https://doi.org/10.3390/su10061700

Wu, W., Kou, G., & Peng, Y. (2018). A consensus facilitation model based on experts' weights for investment strategy selection. *Journal of the Operational Research Society*, *69*(9), 1435–1444. https://doi.org/10.1080/01605682.2017.1398203

~~Xue, R., Wang, C., Liu, M., Zhang, D., Li, K., & Li, N. (2019). A new method for soil health assessment based on analytic hierarchy process and meta-analysis. *Science of the Total Environment*, *650*, 2771–2777. https://doi.org/10.1016/j.scitotenv.2018.10.049~~

~~Zhang, L., Lavagnolo, M. C., Bai, H., Pivato, A., Raga, R., & Yue, D. (2019). Environmental and economic assessment of leachate concentrate treatment technologies using analytic hierarchy process. *Resources, Conservation and Recycling*, *141*, 474–480. https://doi.org/10.1016/j.resconrec.2018.11.007~~

Zhang, X., Xiong, G., Cui, F., Zhao, Q., & Liu, H. (2018). Study on geothermal energy utilization technology evaluation based on AHP and multi-level fuzzy comprehensive evaluation. *IOP Conference Series: Earth and Environmental Science*, *208*(1), 12044. https://doi.org/10.1088/1755-1315/208/1/012044