



Modeling Group Perceptions Using Stochastic Simulation: Scaling Issues in the Multiplicative AHP

Barfod, Michael Bruhn; van den Honert, Robin; Salling, Kim Bang

Published in:

International Journal of Information Technology and Decision Making

Link to article, DOI:

[10.1142/S0219622016500103](https://doi.org/10.1142/S0219622016500103)

Publication date:

2016

Document Version

Peer reviewed version

[Link back to DTU Orbit](#)

Citation (APA):

Barfod, M. B., van den Honert, R., & Salling, K. B. (2016). Modeling Group Perceptions Using Stochastic Simulation: Scaling Issues in the Multiplicative AHP. *International Journal of Information Technology and Decision Making*, 15(2), 453-474. <https://doi.org/10.1142/S0219622016500103>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

MODELLING GROUP PERCEPTIONS USING STOCHASTIC SIMULATION: SCALING ISSUES IN THE MULTIPLICATIVE AHP

MICHAEL BRUHN BARFOD*

*Department of Transport
Technical University of Denmark
Bygningstorvet 115, DK-2800 Kgs. Lyngby, Denmark
mbb@transport.dtu.dk*

ROB VAN DEN HONERT

*Risk Frontiers - Natural Hazards Research Centre
Macquarie University
North Ryde, NSW 2109, Australia
rob.vandenhonert@mq.edu.au*

KIM BANG SALLING

*Department of Transport
Technical University of Denmark
Bygningstorvet 115, DK-2800 Kgs. Lyngby, Denmark
kbs@transport.dtu.dk*

Received (Day Month Year)

Revised (Day Month Year)

Communicated by (xxxxxxx)

This paper proposes a new decision support approach for applying stochastic simulation to the multiplicative AHP in order to deal with issues concerning the scale parameter. The paper suggests a new approach that captures the influence from the scale parameter by making use of probability distributions. Herein, the uncertainty both with regard to the scale and the inherent randomness from the parameter is captured by probabilistic input and output distributions. Provided that each alternative and criteria under consideration are independent it is assumed that the embedded uncertainty from the progression factors remains the same. The result is then an interval estimate for each alternative's final scores. This can lead to overlapping intervals of scores which may be interpreted as possible rank reversals. Thus the decision support approach makes it possible to calculate the probability of overlapping for any given set of pairwise comparisons.

Keywords: Decision Support; Multi-Criteria Decision Analysis; multiplicative AHP; Stochastic Simulation.

1. Introduction

The original Analytic Hierarchy Process (AHP) by Saaty¹, which is based on the additive value function, is a widely applied technique for multi-criteria decision making. The technique makes use of pair wise comparisons to derive scores, and thereby a rank ordering of the alternatives under consideration. The decision-makers are involved in the

* Corresponding author

process by providing subjective cardinal judgments about the preference for each alternative over each of the other alternatives under all criteria. The judgments are assumed to be categorical descriptive judgment values which later on in the process are converted onto a numerical scale. The final set of comparisons is synthesized into scores, making it possible to rank order the considered alternatives.

However, research has criticized the original additive AHP for three reasons: 1) the fundamental scale to quantify human judgments; 2) the estimation of impacts scores by the Perron-Frobenius eigenvector; and 3) the calculation of final scores using an additive model. Zahedi² argued that the criticism of the AHP concentrated on the estimation of the impact scores, but that no major controversy existed concerning the aggregation step. Zahedi² did not mention any criticism of the fundamental scale, but Belton³ brought forward several arguments against the scale and the aggregation rule. Stewart⁴ discussed the three issues above and warned that the additive AHP, despite its widespread popularity, should only be used with considerable caution. Leskinen and Kangas⁵ discussed the connections between rank reversals and the potential inconsistency of the utility assessments in the case of ratio-scale pairwise comparison data, and they concluded that using the geometric mean aggregation instead of the eigenvector aggregation does not cause undesired rank reversals. Korhonen and Topdagi⁶ have also contributed to the discussion and have brought forward arguments regarding the inappropriateness of the ratio scale in certain decision situations. Barzilai et al.⁷, Barzilai and Golany⁸ and Barzilai⁹ made the observation that the additive AHP, since it is initially based upon ratio information, should be converted into a variant with a multiplicative structure. This version of AHP is normally referred to as the multiplicative AHP.

The multiplicative AHP has – just as the additive version – proven to be well suited for group decision making, and several software extensions have been developed for the technique to cope with this. In the simplest form Van den Honert and Lootsma¹⁰ proposed to use a geometric mean group aggregation procedure to satisfy the social choice axioms. A more sophisticated preference model by Barzilai and Lootsma¹¹ includes the group members as a separate level of the decision hierarchy, which means that each of the group members perform their own set of pair wise comparisons under each of the criteria. The final impact scores are then aggregated across the group members using the geometric mean, and hereby point estimates are obtained that reflect the group's common viewpoint. However, the result does not reflect to what degree of consensus has been achieved in the group regarding the comparative judgments.

Van den Honert¹² on the other hand examines the variation in the decision-makers' comparative judgments in a consensus seeking group and proposes a group preference model that expresses the group's preference intensity judgments as random variables with associated probability distributions (in this case, normal distribution). The judgments are then used to determine the probability of rank reversal occurring amongst the alternatives due to the group preference uncertainty.

Jalao et al.¹³ treats the issue of modelling decision makers' varying stochastic (imprecise) preferences by using a beta distribution to better fit the variety in the

preferences. Dede et al.¹⁴ addressed the impact of various aspects that may affect the estimation of probability for rank reversals in the context of pairwise comparisons such as uncertainty level, alternative preference scales and different weight approximation methods. Also Eskandari and Rabello¹⁵ presented a stochastic approach to capture the uncertain behaviour of the global AHP weights. Jato-Espino et al.¹⁶ instead converted the group judgments into triangular fuzzy numbers to capture the vagueness that human attitude entails when making judgments.

The above-mentioned models, however, do not address the decision-makers' perception of the scale used for the comparative judgments. The multiplicative AHP makes use of direct rating on a geometric scale, and for this purpose the progression factor 2 is used for calculating scores of alternatives and $\sqrt{2}$ for calculation of criteria weights when transforming the comparative judgments. Barfod and Leleur¹⁷ addressed this issue by analyzing the sensitivity of the final outcome of an analysis using different progression factors within an interval. However, the work assumed consensus about the progression factor in the group of decision-makers. Depending on the decision context, the decision-makers' aversion towards risk, etc., it is, however, most likely that the decision-makers have varying progression factors in their perception of the scale. This potential variability in the progression factors implies a form of uncertainty in the technique which may be modelled by stochastic simulation. Salling¹⁸ examined the use of stochastic simulation for the detection of uncertainties in quantitative decision support techniques, and Rodger et al.¹⁹ examined the uncertainty related to linguistic terms in the weighted averaging approach using stochastic simulation and disclosed its usefulness in group decision making. Hence the techniques seem appropriate to use for the current research issue. Using stochastic simulation interval results will be obtained, and the probability of rank reversal between the alternatives caused by the group's different perceptions of the scale can be calculated.

This paper does not attempt to contribute to the discussion of the appropriateness of the additive versus the multiplicative AHP – this has already been extensively covered by the literature (see e.g. Refs. 2–4, 6, 11 and 20 for more details). Instead the paper only focuses on one specific scaling issue in the multiplicative version of AHP, namely, the uncertainty related to the progression factors used when transforming the decision-makers' comparative judgments from the verbal to the geometric scale. The influence of and the uncertainty associated with the progression factors are examined by addressing different possible perceptions of the scale parameters using stochastic simulation. Thus the paper presents a new approach in the form of a decision support technique which is able to model the uncertainty relating to the progression factors.

After this introduction Sec. 2 will describe the multiplicative AHP and elaborate on the mathematics used for the purpose of this paper. Sec. 3 outlines the stochastic model applied and presents the results of the test simulations on three different case studies. The results are discussed in Sec. 4, and Sec. 5 presents the conclusions.

2. The multiplicative AHP

As for the original (additive) AHP, the multiplicative version makes use of a structured hierarchical approach based on the principle that decision-makers make pair wise comparisons between alternatives to determine subjective impacts under each criterion in the assessment, and between criteria in order to determine their relative importance. Finally, aggregating the results leads to a final score for each project, which allows a subjective rank ordering of the alternatives.^{20,21,22}

The systematic pair wise comparison approach is one of the cornerstones of the multiplicative AHP by Lootsma.²¹ The technique makes use of a procedure for direct rating which requires the decision-makers to consider all possible pairs of alternatives with respect to each criterion in turn in order to determine which one of the alternatives in the pair is preferred, and to specify the strength of preference according to a verbal category scale (with an associated numeric 0-8 scale). The approach is, as mentioned, a multiplicative development of the AHP, and it proposes to overcome the three well-known issues regarding the theory behind AHP.

First, the direct rating is on a geometric scale²¹ which replaces Saaty's 1 – 9 original scale. Second, the eigenvector method originally used in AHP is replaced by the geometric mean method, which avoids potential rank reversal.^{6,23} Third, the aggregation of scores using the additive model is replaced by a multiplicative structure.²⁴

2.1. Pairwise comparisons and scaling

In the use of the multiplicative AHP in this paper it is assumed that the ratifying group consists of g decision-makers ($g \geq 1$), and that at any stage of the process there are n alternatives ($n \geq 1$) under consideration. At the first evaluation level of the analysis, two stimuli S_j and S_k (two alternatives A_j and A_k under a criterion) are presented to the decision-makers. The decision-makers are then asked to express their graded comparative judgment about them. That is, the decision-makers express their indifference between the two, or a weak, definite, strong or very strong preference for one stimulus over the other (see Table 1). Thus, at this stage the decision-makers are asked to make as standard $n(n-1)/2$ pair wise comparisons. Indeed only $(n-1)$ properly chosen comparisons would be sufficient, for which reason the standard leads to much more information being collected than actually needed²⁵. Such redundancy, however, is usually beneficial as it enables a smoothing of the results of the analysis.

It is assumed that the two stimuli have unknown subjective values V_j and V_k . The decision-makers' pair wise comparative judgment of S_j versus S_k is captured on a category scale to frame the range of possible verbal responses. This is converted into an integer-valued gradation index δ_{jk} according to the scale in Table 1. The number of categories is rather small as human beings' linguistic capacity to describe the categories unambiguously in verbal terms is limited.²⁶

Table 1. The multiplicative AHP scale

| Comparative judgment | Gradation index δ_{jk} |
|---|-------------------------------|
| Very strong preference for S_k over S_j | -8 |
| Strong preference for S_k over S_j | -6 |
| Definite preference for S_k over S_j | -4 |
| Weak preference for S_k over S_j | -2 |
| Indifference | 0 |
| Weak preference for S_j over S_k | +2 |
| Definite preference for S_j over S_k | +4 |
| Strong preference for S_j over S_k | +6 |
| Very strong preference for S_j over S_k | +8 |

Intermediate integer values can be assigned to δ_{jk} to express a hesitation between two adjacent categories. The gradation index δ_{jk} can be converted into a value on a geometric scale, characterised by a scale parameter $\gamma = \ln(I+\varepsilon)$, where $I+\varepsilon$ is the progression factor. Thus

$$r_{jk} = \exp(\gamma \delta_{jk}) \quad (1)$$

is defined to be the numeric estimate of the preference ratio V_j/V_k . Although there is no unique scale of human judgment, a plausible value of γ is $\ln(2)$ implying a geometric scale with the progression factor 2.

There are five major, linguistically distinct categories in Table 1: indifference, weak, definite, strong and very strong. Moreover, there are four so-called threshold categories between them which can be used if the decision-makers are in-between the neighbouring qualifications. Lootsma²⁶ shows that human beings follow the same pattern in many unrelated areas when they categorise an interval, e.g. certain ranges on the time axis, and sound and light intensities. Normally three to five major categories are introduced, and the progression factor $\exp(2\gamma) = (I+\varepsilon)2$ is roughly 4 (See Refs. 21 or 26 for more details). By the interpolation of threshold categories a more refined subdivision of the given interval is obtained. In that case there are six to nine categories and the progression factor $\exp(\gamma) = (I+\varepsilon)$ is roughly 2 ($\gamma = \ln 2 \sim 0.7$), which defines what Lootsma²⁷ calls the ‘natural’ scale. In addition, Lootsma²⁷ suggests that sensitivity analysis should be carried out with a ‘short’ ($\gamma = 0.5$) and a ‘long’ ($\gamma = 1.0$) geometric scale in the neighbourhood of the natural scale.

2.2. Pair wise comparisons of alternatives

We now assume that a decision maker compares alternatives under a series of criteria. Individual judgments when comparing alternative j to alternative k under criterion i can be written as δ_{ijk} . Assume that Γ is a scale parameter, which has an element of uncertainty with respect to scale and inherent randomness (we call these uncertainties ‘unknownness’). We then assume that each group members’ perception of the chosen value of δ_{ijk} can be modelled by a probability distribution, which variance is a measure of the lack of unanimity in the group. We can then model the ‘unknownness’ in Γ by a probability distribution. Then γ is a random variable, and so

$$r_{ijk} = \exp(\gamma \delta_{ijk}) \quad (2)$$

is the numerical scale of the verbal comparative judgments. Since γ is a random variable, so is r_{ijk} . Now, using the standard multiplicative model, the final scores for the alternatives are

$$S_j = \prod_{i=1}^m \left(\prod_{k=1}^n r_{ijk}^{1/n} \right)^{c_i} \quad (3)$$

where m is the number of criteria, n is the number of alternatives, and c_i is the normalised criteria weight. So

$$\begin{aligned} \ln S_j &= \frac{1}{n} \sum_{i=1}^m \sum_{k=1}^n c_i r_{ijk} \\ &= \frac{1}{n} \sum_{i=1}^m \sum_{k=1}^n c_i \gamma \delta_{ijk} \\ &= \frac{\gamma}{n} \sum_{i=1}^m \sum_{k=1}^n c_i \delta_{ijk} \end{aligned}$$

It is assumed that $\gamma \sim N(\mu_\gamma, \sigma_\gamma^2)$. This is a plausible assumption as the value for γ is likely to be symmetrically distributed around the mean value, with the probability of extreme values on both sides of this being small.¹² For the alternatives we might assume $\mu_\gamma = \ln 2$.

$$\ln S_j \sim N(\mu_{\ln S_j}, \sigma_{\ln S_j}^2)$$

with mean

$$\mu_{\ln S_j} = \frac{\mu_\gamma}{n} \sum_{i=1}^m \sum_{k=1}^n c_i \delta_{ijk} \quad (4)$$

and variance

$$\sigma_{\ln S_j}^2 = \frac{\sigma_\gamma^2}{n^2} \left(\sum_{i=1}^m \sum_{k=1}^n c_i^2 \delta_{ijk}^2 + 2 \sum_{i=1}^m c_i^2 \sum_{k < \rho} \sigma_{\delta_{ijk}, \delta_{ij\rho}} \right) \quad (5)$$

The second term within parentheses in Eq. (5) falls away if we assume the components of the impact score vector are independent.

Since the $\ln S_j$ is approximately normally distributed, the S_j are then asymptotically lognormal distributed, with the approximation improving as m and n get larger, and as the distribution of γ tends to normality. Assuming m and n are not “too” small, then mn will be large enough to ensure reasonable normality of the $\ln S_j$, and lognormality of the S_j . Then

$$S_j \in \exp\left(\mu_{\ln S_j} \pm z_{\alpha/2} \sqrt{\sigma_{\ln S_j}^2}\right), \quad j = 1 \dots n$$

In most practical applications both m and n can be expected to be rather small (e.g. $m, n \leq 10$), but the product mn might be rather large, which implies a reasonable approximation of normality. Generally, if $mn > 30$ the sample can be considered large enough for ensuring normality.

It should be noted that m in most cases will be larger than 10 during the initial phases of an appraisal study, but for the final phase of decision making the criteria is often scrutinized and reduced to a maximum of 8-10. This as criteria with an importance ranking lower than this number do not contribute significantly to the differentiation between alternatives, i.e. they are without actual influence on the decision. Thus in a practical application m is often less than 10, but it is most likely larger than 10 in more theoretical applications. This paper considers the practical use; hence the assumption above.

2.3. Pair wise comparisons of criteria

Comparisons between criteria work exactly the same way, but makes use of a different scale parameter value²⁷, $\gamma = \ln\sqrt{2}$. However if we introduce uncertainty here then there is uncertainty in two separate places in the model, i.e. in both γ and c_i .

Now we have

$$\ln S_j = \frac{\gamma}{n} \sum_{i=1}^m \sum_{k=1}^n c_i \delta_{ijk}$$

where

$$\gamma \sim N(\mu_\gamma, \delta_\gamma^2)$$

and

$$c_j = \left(\prod_{k=1}^n r_{jk}\right)^{1/n} = \left(\prod_{k=1}^n \exp(\gamma \delta_{jk})\right)^{1/n} \quad (6)$$

where $r_{jk} = \exp(\gamma \delta_{jk})$. Then

$$\ln c_j = \frac{\gamma}{n} \sum_{k=1}^n \delta_{jk}$$

If we assume again that

$$\gamma \sim N(\mu_\gamma, \sigma_\gamma^2)$$

then

$$\ln c_j \sim N(\mu_{\ln v_j}, \sigma_{\ln v_j}^2)$$

where

$$\mu_{\ln v_j} = \frac{\mu_\gamma}{n} \sum_{k=1}^n \delta_{jk} \quad (7)$$

and

$$\sigma_{lnv_j}^2 = \frac{\sigma_v^2}{n^2} \left[\sum_{k=1}^n \delta_{jk}^2 + 2 \sum_{k < p} \sigma_{\delta_{jk}, \delta_{jp}} \right] \quad (8)$$

Again the second term in parentheses in Eq. (8) falls away if the components of the impact score vector are independent.

2.4. Stability of the (un-normalised) impact score vector

The probability that alternative i will reverse rank with some other alternative due to the lack of unanimity in the estimate of the scale used for comparative judgements is given by

$$P_i = 1 - \prod_{j=1}^n (1 - P_{ij}), \quad i = 1 \dots n \quad (9)$$

where

$$P_{ij} = Pr(IV_{jl} \leq \ln S_i \leq IV_{iu}) Pr(IV_{jl} \leq \ln S_j \leq IV_{iu})$$

and the probability of at least one rank reversal occurring in the system due to group scaling uncertainty is given by

$$P = 1 - \prod_{1 \leq i < j \leq n} (1 - P_{ij}) \quad (10)$$

where P_{ij} are defined as above. This is depicted in Figure 1.

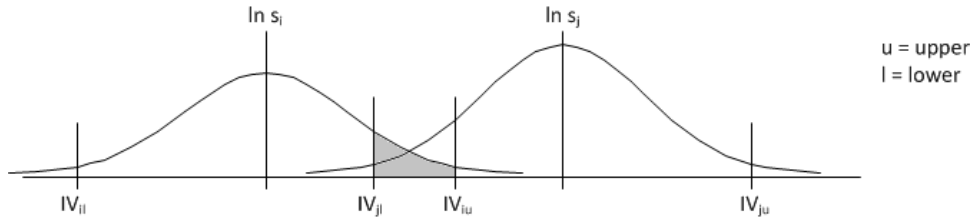


Figure 1. Overlapping intervals

The quantity P is a measure of the overall instability of the vector of un-normalised impact scores given the lack of unanimity in the estimate of the scale used for comparative judgements.

3. Stochastic simulations

The uncertainty in the progression factors is considered using stochastic simulation. In this respect the calculations of the uncertainties related to the preference comparisons are conducted using Monte Carlo Simulation (MCS), which is a computerized mathematical technique that simulates the uncertainties embedded in most quantitative analyses and decision making. It provides the decision-maker with a range of possible outcomes and

associated probabilities for any choice of action. MCS thus performs the analysis by building models of possible results by substituting a range of values – a probability distribution – for any parameter or variable that has inherent uncertainty. It then reproduces the results e.g. scores of an alternative, over and over, each time using a different set of random values from the probability distribution. By using probability distributions for the simulations, the probabilities of different outcomes occurring can be determined, as depicted in Figure 1.

During a MCS, values are sampled at random from the input probability distributions. Each set of samples is called an iteration, and the resulting outcome from that sample is recorded. MCS does this hundreds or thousands of times, and the result is a probability distribution of possible outcomes. It has previously been found that the Latin Hypercube sampling technique satisfies the MCS process by recreating the input distribution through a stratified sampling without replacement method.¹⁸ This means that the input probability distributions are divided into intervals on the cumulative curve. The sampling procedure is then forced to represent the values in each interval, thus creating the input distribution. Further advantages, disadvantages and pitfalls of simulation are discussed in details by Law and Kelton.²⁸ The choice of probability distribution to model the issue of this paper is elaborated below.

3.1. Probability distribution

Typically, any assessment scheme relies on available data to represent, for example, traffic or passenger flows. Literature and current research has proven that parametric distributions such as the Normal, Gamma or Beta distributions are able to capture and hence fit data from such flow models.^{29,30} The choice and validity of the selected distribution are based upon expert opinions since data with respect to the progression factor are sparse. Vose³¹ suggest the use of either a Triangular distribution or a beta-PERT distribution, hereinafter referred to as a PERT distribution, where input is based upon a minimum, maximum and most likely value.

The PERT distribution is derived from the beta distribution, which covers a huge variety of skewness types. When used in MCS, the PERT distribution can be used to identify uncertainties based on its similarity to the triangular distribution. Just as the triangular distribution, the PERT distribution sets emphasis on the most likely value over the minimum and maximum estimates, contributing to the non-parametric dimension of the distribution function. However, in contradiction to the triangular distribution the PERT distribution constructs a smooth curve, which places more emphasis on values close to the most likely value.

As discussed by Back et al.³² and Salling and Leleur³³ the beta-PERT distribution becomes attractive when the specific uncertain parameter/value is categorised possessing:

- (1) Upper and lower limits which ensures that the analyst is relatively certain the parameter/value does not exceed; consequently, a closed-ended distribution is desirable;
- (2) The distribution must be continuous;
- (3) The distribution will be unimodal, presenting a most likely value, and
- (4) The distribution must be able to have a greater freedom to be

higher than lower with respect to the estimation – skewness must be expected. Thus, as illustrated in Table 2-4, all four items can be properly addressed.

In real-life decision situations people are most often able to give a more confident estimate of the most likely value, rather than the extreme values. In this paper we assume that the same situation is the case with people's perception of the progression factor in the multiplicative AHP, where the most likely values are set to 2 and $\sqrt{2}$ for alternatives and criteria respectively, in accordance with Lootsma's 'natural' scale. For the minimum and maximum values we use 1.7 and 2.7 respectively – the 'short' and 'long' scales which were recommended by Lootsma²⁷ and confirmed by Barford and Leleur.¹⁷ Hence, the PERT distribution brings a smooth description of the tail of the factors considered. In practice, this implies that we trust the most likely estimate, and we believe that even though it is not accurate, we still expect the resulting values to be close to that estimate.

3.2. Case studies

Simulations have been conducted on three different case studies in order to reveal possible general tendencies with regard to the alternatives' behaviour under the criteria when uncertainty is introduced to the progression factors. The case studies have in common that they are all transport-related, and are assessed using the same procedure: Using the concept of decision conferencing³⁴, input has been generated to the technique by decision makers and stakeholders. In this paper the case studies are only used as illustrative examples of the simulations, and the results have therefore not been presented to the respective groups that generated the input.

The cases have been examined in three different scenarios: one where only the progression factor for the alternatives is addressed, one where only the progression factor for the criteria is addressed, and finally one where both progression factors are addressed simultaneously.

3.2.1. The Rail Baltica case

The case study concerns alternatives for a new high speed railway link connecting Helsinki (Finland) – Tallinn (Estonia) – Riga (Latvia) – Kaunas (Lithuania) – Warsaw (Poland) and continuing to Berlin (Germany). This case study concerns the first step in connecting Helsinki and Berlin, i.e. the link going from Tallinn through the Baltic countries: Estonia, Latvia and Lithuania, to the border between Lithuania and Poland. The case, which is described in detail by Ambrasaite et al.³⁵ and further treated by Petursson³⁶, considers four alternatives for upgrading the railway namely: the shortest route consisting of mostly new alignment (A1), the most direct route consisting of existing alignment (A2), the route with the largest potential passenger demand consisting of a combination of new and existing alignment (A3), and the route that utilizes all existing routes along the corridor (A4). The alternatives were assessed under nine criteria.

One of the criteria makes use of input from a preliminary socio-economic assessment of the alternatives. This socio-economic assessment included impacts such as construction costs, operating costs, travel time savings, ticket revenue and emissions. The

results in terms of net present values for the alternatives were then used as a basis for the comparisons under the criterion. The resulting scores obtained from the preliminary calculations before uncertainty is introduced in the technique were: 0.52 (A1), 0.13 (A2), 0.20 (A3), and 0.15 (A4).

A1 (with a total score on 0.52) is clearly preferable to the three other alternatives. A2, A3 and A4 on the other hand obtain scores that are very close to each other. Applying probability distributions to the progression factors will reveal if any overlapping score-intervals (implying rank reversal issues) will occur between the alternatives.

The results of the stochastic simulations are shown on a graph, which has the probability density on the y-axis and the normalised total score on the x-axis. Hence it is possible to read the probability of a possible outcome of the calculations on this basis. The maximum density (near the middle of the graphs) depicts the most likely total score while the tails depict the minimum (min) and maximum (max) scores which are the least likely to occur.

Figure 2 depicts the results of the first simulations conducted on the progression factor for the alternatives only. As it can be seen from Figure 2 the choice of A1 as the most preferred seems very robust as none of the other alternatives' score-intervals come even close to this. However, it can be seen that great uncertainty is associated with identifying which alternative should be ranked as number two, three and four. The mean total scores for A2, A3 and A4 in Table 2 (depicted by the vertical line on the figure) were very close to each other, and the simulations reveal that in such a case different perceptions of the progression factor can lead to disagreements in the group of decision-makers.

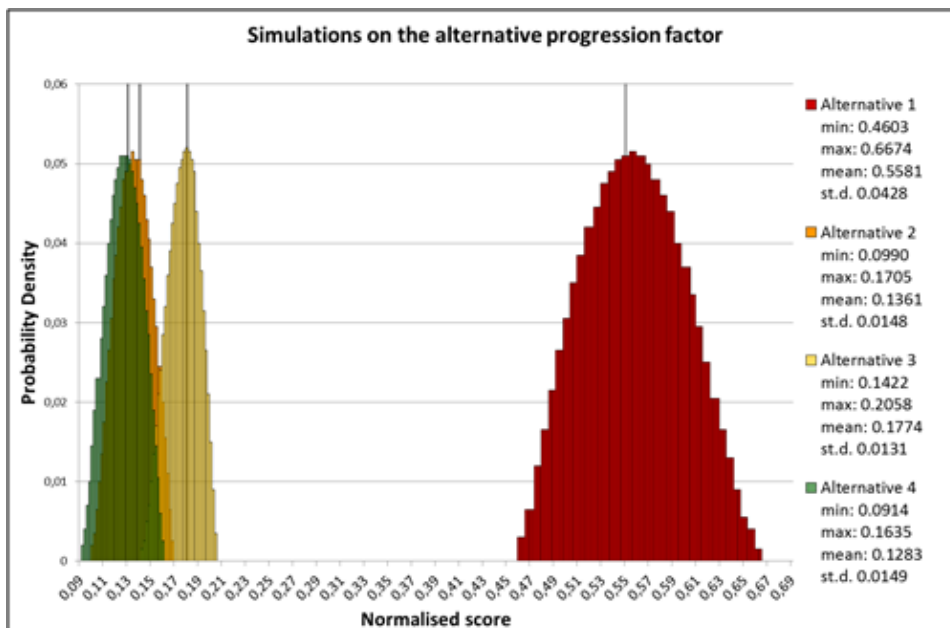


Figure 2. Simulations on the alternative progression factor: Rail Baltica case

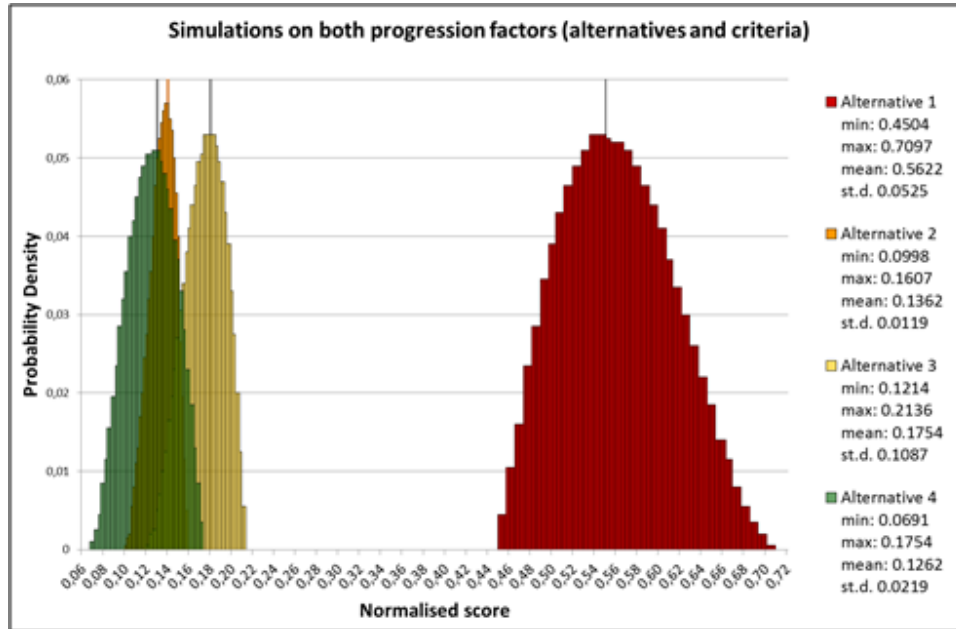


Figure 3. Simulations on both the alternative and criteria progression factor: Rail Baltica case

Exactly the same tendency is shown when simulating the progression factor for criteria only, and when simulating both progression factors simultaneously. Figure 3, which contains simulations on both progression factors, depicts that the intervals between min and max values have become marginally wider in the latter case, implying a greater uncertainty.

The different simulations show that no dispute exists about which alternative is the most preferable for the Rail Baltica case. However, great uncertainty is related to the ranking of the remaining three alternatives, all of which could be the second most preferred, the third most preferred, and the least preferred if the uncertainty in the progression factor is considered. If the decision making process requires such a ranking, further examinations and assessments need to be made that can assist to the segregation of the alternatives.

Table 2 summarises the results of the simulations in the three different scenarios for the case study. The mean values (mean), standard deviations (st.d.) and probabilities are calculated in accordance with respectively Eq. (4), Eq. (5), Eq. (7), Eq. (8), Eq. (9), and Eq. (10) in Sec. 2. The P_i values indicate each alternative's probability for reversing rank with another alternative, while the P values is the probability of the system experiencing at least one rank reversal. As seen in Table 2 the latter is 1.000 in the first and the last scenarios indicating that a rank reversal is certain to take place.

Table 2. Results from the simulations on the Rail Baltica case study

| Scenario | | A1 | A2 | A3 | A4 |
|---|----------------------------|-------|-------|-------|-------|
| Alternative progression factor (1.7; 2; 2.7) | Min. | 0.460 | 0.099 | 0.142 | 0.091 |
| | Max. | 0.667 | 0.171 | 0.206 | 0.164 |
| | μ_{ins_i} (mean) | 0.558 | 0.136 | 0.177 | 0.128 |
| | $\sigma_{ins_i}^2$ (st.d.) | 0.043 | 0.015 | 0.013 | 0.015 |
| | P_i | 0.000 | 1.000 | 0.360 | 0.976 |
| | P (system) | | | | |
| Criteria progression factor ($\sqrt{1.7}$; $\sqrt{2}$; $\sqrt{2.7}$) | Min. | 0.532 | 0.121 | 0.166 | 0.115 |
| | Max. | 0.578 | 0.144 | 0.192 | 0.146 |
| | μ_{inv_i} (mean) | 0.550 | 0.140 | 0.180 | 0.130 |
| | $\sigma_{inv_i}^2$ (st.d.) | 0.009 | 0.003 | 0.005 | 0.007 |
| | P_i | 0.000 | 0.475 | 0.000 | 0.525 |
| | P (system) | | | | |
| Both progression factors | Min. | 0.450 | 0.078 | 0.121 | 0.069 |
| | Max. | 0.710 | 0.182 | 0.214 | 0.175 |
| | μ_{inv_i} (mean) | 0.562 | 0.136 | 0.175 | 0.126 |
| | $\sigma_{inv_i}^2$ (st.d.) | 0.053 | 0.034 | 0.109 | 0.022 |
| | P_i | 0.000 | 1.000 | 0.418 | 0.869 |
| | P (system) | | | | |

3.2.2. The HH-case

The Oresund fixed link connecting the area of greater Copenhagen with Malmö in Sweden opened in July 2000. Today, fifteen years later, the railway line of the link is close to its capacity limit, resulting in delays and discomfort for the travellers. The case concerns a new complementary fixed link over Oresund (the sea separating the island of Zealand (Denmark) with Sweden) between the cities of Elsinore (Helsingor) and Helsingborg and is normally referred to as the HH-connection. Regionally, the proposed connection is expected to create a substantial increase in trade, education and work-related benefits. Ultimately it is expected that a fixed link with increased commuter traffic across the border will result in a common labour and residence market. In addition, the recent decision to construct the Fehmarn Belt fixed link between Denmark and Germany will increase the number of travellers from central Europe through Denmark to Sweden, Norway and Finland. This means additional traffic to cross the Oresund. The case, which is described in details Barfod and Leleur¹⁷ and Barfod and Salling³⁷, considers four alternatives, namely: a railway tunnel for passengers only (A1), a railway tunnel for goods only (A2), a bridge for both road and rail (A3), and a bridge only for road (A4). The alternatives were assessed under five criteria using the multiplicative AHP.

The resulting scores obtained from the preliminary calculations before uncertainty is introduced in the technique were: 0.12 (A1), 0.01 (A2), 0.47 (A3), and 0.40 (A4). A3 (with a total score of 0.47) is preferable to the three other alternatives. However, A4 follows closely with a total score of 0.40, while A1 and especially A2 are inferior in this context. Applying probability distributions to the progression factors will reveal if any overlapping score-intervals will occur between the two most preferred alternatives.

Figure 4 depicts the results of the first simulations conducted on the progression factor for the alternatives only. As it can be seen from the figure the choice of A3 as the most preferred seems robust as none of the other alternatives' score intervals come close to this. In this scenario the ranking of all the alternatives can be identified without uncertainty as none of the intervals of scores overlap with each other. Hence different perceptions of the progression factor for alternatives do not lead to disagreements regarding the ranking of alternatives in the group of decision-makers.

Simulations conducted on the progression factor for the criteria show the same tendency as in Figure 4. However, if we simulate on both progression factors simultaneously an overlap between the score intervals for A3 and A4 occur, cf. Figure 5. This implies that decision-makers with different perceptions of the progression factors can be in conflict about which alternative is the most preferred.

In the situation depicted in Figure 5 it is relevant to determine the probability (P_i) of rank reversal for each of the conflicting alternatives. The results are summarised in Table 3 where it can be seen that no rank reversals will occur in the first two scenarios (the P values are 0). In the third scenario A3 has a P_i value of 0.043, which indicates that the probability of rank reversal with other alternatives is very low. The overall P value is 0.221 indicating a low probability for the issue to arise in the system. Conflicts will thus only occur if the most extreme values of the input intervals are used. However, in a situation where the original scores of the two (or more) most preferred alternatives are closer than here, the uncertainty can be assumed to be much higher. The simulations show no disputes about the ranking of A1 and A2, which are clearly dominated by A3 and A4.

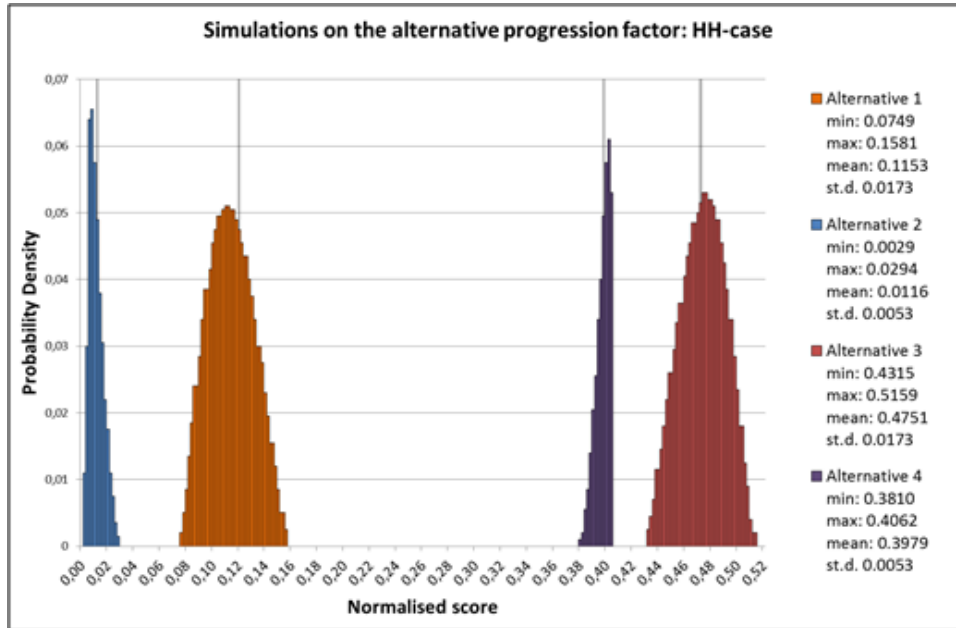


Figure 4. Simulations on the alternative progression factor: HH-case

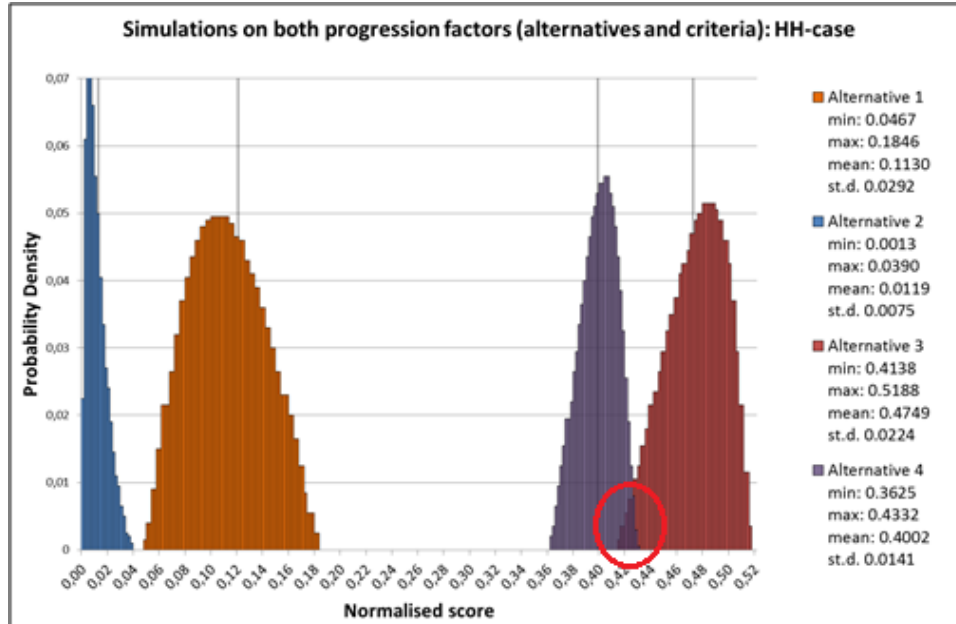


Figure 5. Simulations on both the alternative and criteria progression factor: HH-case

Table 3. Results from the simulations on the HH case study

| Scenario | | A1 | A2 | A3 | A4 |
|---|----------------------------|-------|-------|-------|-------|
| Alternative progression factor (1.7; 2; 2.7) | Min. | 0.075 | 0.003 | 0.432 | 0.381 |
| | Max. | 0.158 | 0.029 | 0.516 | 0.406 |
| | μ_{ins_i} (mean) | 0.115 | 0.012 | 0.475 | 0.398 |
| | $\sigma_{ins_i}^2$ (st.d.) | 0.017 | 0.005 | 0.017 | 0.005 |
| | P_i | 0.000 | 0.000 | 0.000 | 0.000 |
| | P (system) | | | | |
| Criteria progression factor ($\sqrt{1.7}$; $\sqrt{2}$; $\sqrt{2.7}$) | Min. | 0.088 | 0.007 | 0.454 | 0.391 |
| | Max. | 0.150 | 0.018 | 0.481 | 0.402 |
| | μ_{inv_i} (mean) | 0.117 | 0.012 | 0.472 | 0.399 |
| | $\sigma_{inv_i}^2$ (st.d.) | 0.013 | 0.002 | 0.006 | 0.003 |
| | P_i | 0.000 | 0.000 | 0.000 | 0.000 |
| | P (system) | | | | |
| Both progression factors | Min. | 0.047 | 0.001 | 0.414 | 0.363 |
| | Max. | 0.185 | 0.039 | 0.519 | 0.433 |
| | μ_{inv_i} (mean) | 0.113 | 0.012 | 0.475 | 0.400 |
| | $\sigma_{inv_i}^2$ (st.d.) | 0.029 | 0.008 | 0.022 | 0.014 |
| | P_i | 0.000 | 0.000 | 0.043 | 0.186 |
| | P (system) | | | | |

3.2.3. The Ostlänken case

The case study concerns an assessment of alternatives for a new high-speed railway line in Sweden named Ostlänken. More specifically, the case study considers a section between Bäckeby and Norrköping about 100 kilometres south of Stockholm. The case, which is described in detail by Barfod et al.³⁸, considers four alternatives, namely: an upgrade of the existing railway (A1), a new railway line with a short tunnel under a part of a natural habitat (A2), a new railway line with a long tunnel under a natural habitat (A3), and a new railway line around a natural habitat (A4). The alternatives were assessed under nine criteria using the multiplicative AHP.

The resulting scores obtained from the preliminary calculations before uncertainty is introduced in the technique were: 0.18 (A1), 0.38 (A2), 0.32 (A3), and 0.12 (A4). A2 (with a total score of 0.38) is preferable to the three other alternatives. However A3 follows closely with a total score of 0.32, while A1 and A4 are dominated. Applying probability distributions to the progression factors will reveal if any overlapping score-intervals will occur between the two most preferred alternatives.

Figure 6 depicts the results of the simulations conducted on the progression factor for the alternatives only. As can be seen from the figure the choice of A2 as the most preferred is to some degree uncertain as the score interval overlaps with the interval of A3. This again implies that decision-makers with different perceptions of the progression factors can be in conflict about which alternative is the most preferred. The probability (P_i) of A2 experiencing rank reversal with another alternative is in this scenario 0.077. A conflict also occurs between A1 and A4, and the probabilities for rank reversals are 0.041 and 0.039 respectively.

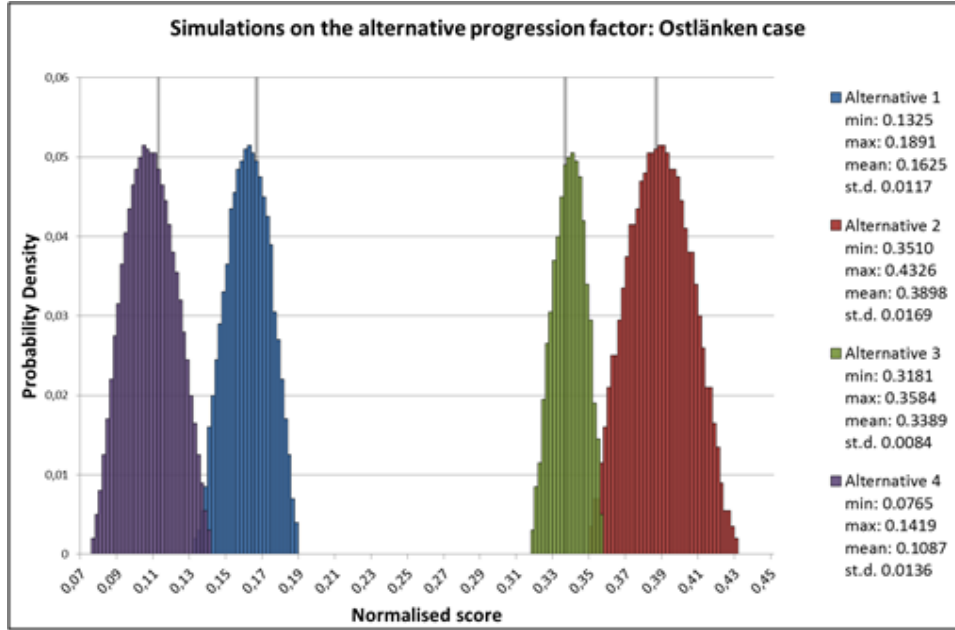


Figure 6. Simulations on the alternative progression factor: Ostlänken case

The overlaps also occur when simulating on the criteria progression factor and when simulation on both progression factors simultaneously. Table 4 summarises the results of the simulations in the three different scenarios, and it can be seen that the overall probabilities for rank reversals to take place in the system are relatively low.

Table 4. Results from the simulations on the Ostlänken case study

| Simulation | | A1 | A2 | A3 | A4 |
|---|----------------------------|-------|-------|-------|-------|
| Alternative progression factor (1.7; 2; 2.7) | Min. | 0.143 | 0.351 | 0.318 | 0.077 |
| | Max. | 0.189 | 0.433 | 0.359 | 0.142 |
| | μ_{ins_i} (mean) | 0.163 | 0.389 | 0.332 | 0.109 |
| | $\sigma_{ins_i}^2$ (st.d.) | 0.016 | 0.017 | 0.011 | 0.012 |
| | P_i | 0.041 | 0.077 | 0.022 | 0.039 |
| | P (system) | | | | |
| Criteria progression factor ($\sqrt{1.7}$; $\sqrt{2}$; $\sqrt{2.7}$) | Min. | 0.157 | 0.360 | 0.322 | 0.081 |
| | Max. | 0.186 | 0.395 | 0.358 | 0.120 |
| | μ_{inv_i} (mean) | 0.173 | 0.383 | 0.338 | 0.113 |
| | $\sigma_{inv_i}^2$ (st.d.) | 0.011 | 0.007 | 0.008 | 0.002 |
| | P_i | 0.000 | 0.033 | 0.026 | 0.000 |
| | P (system) | | | | |
| Both progression factors | Min. | 0.137 | 0.349 | 0.284 | 0.067 |
| | Max. | 0.222 | 0.443 | 0.367 | 0.143 |
| | μ_{inv_i} (mean) | 0.172 | 0.387 | 0.332 | 0.110 |
| | $\sigma_{inv_i}^2$ (st.d.) | 0.019 | 0.018 | 0.017 | 0.014 |
| | P_i | 0.073 | 0.107 | 0.053 | 0.066 |
| | P (system) | | | | |

4. Discussion

The methodology developed in this paper provides an alternative approach to examining possible differences in scale perceptions in a consensus seeking group. Previous research has been focussed on the situation where decision makers disagree on the judgments (see Refs. 11-13 and 16 for more details) and potential rank reversals due to the aggregation rules (see Refs. 5, 7-9 and 11 for more details). The approach presented in this paper has focussed on a single scale issue in the multiplicative AHP, namely, the uncertainty related to the progression factors defining the geometric scale.

After going through the proposed aggregation procedure, the intervals for the final scores of the alternatives can be used to synthesise an indication of the overall degree of consensus within the group about the final rank order of the alternatives. It is proposed that the groups' overall degree of consensus can be calculated as the probability of no possible rank reversals in the system. The closer this overall P value is to 0, the higher degree of consensus exists in the group. Exactly what P should be to declare sufficient consensus is a subjective judgment which needs to be made on the basis of the composition of the group, the purpose of the analysis, the importance, and other such factors.

The choice - or setting - of the scale parameter in the multiplicative AHP is rather arbitrary and the natural scale is determined by examining peoples' general perception of the progression between steps on different category scales (see Ref. 26 for more details). Thus it is most likely that different perceptions on the parameter exist in a group. This model uncertainty is addressed using stochastic simulations conducted within an interval that has been defined by previous research by Barfod and Leleur¹⁷ and Lootsma²⁷. The interval makes use of a short, a natural, and a long scale, which have been found appropriate for sensitivity tests. It can, however, be assumed that there is a connection between the knowledge level of the group members using the scale and the level of uncertainty. A technical/experienced user may be more likely to use the scale correctly (use the natural scale) than a user without any specific technical knowledge who may apply the scale for the first time. This issue is dealt with by using the beta-PERT distribution for the Monte Carlo simulation within the interval as it is assumed that the natural scale parameter is the most likely and most emphasis should be on this.

Three cases studies have been examined to test the significance of the results of the simulations in situations where the original weight scores of the alternatives are distributed in non-similar ways. Three very similar (all transport related) studies were chosen – not to demonstrate different effects of the proposed methodology – but instead to reveal potential similarities in the results from the approach. Clearly more case studies would need to be examined to make any final conclusions on this, but the three can provide an initial indication of the significance of the issue.

The first case study concerning the Rail Baltica shows four alternatives where one alternative (A1) strongly dominates the others (0.52 (A1), 0.13 (A2), 0.20 (A3), and 0.15 (A4)). Here it is clear that the ranking of the dominated alternatives (A2, A3 and A4) are very sensitive towards changes in the scale parameter. The second case study concerning

the HH-connection shows four alternatives with a clear ranking seen from the original weight scores (0.12 (A1), 0.01 (A2), 0.47 (A3), and 0.40 (A4)). The simulations reveal that the introduced uncertainty in the model is strong enough to result in a rank reversal between A3 and A4. In the third and final case study concerning the Ostlänken railway corridor the four alternatives have original scores that are more close to each other than the two other cases (0.18 (A1), 0.38 (A2), 0.32 (A3), and 0.12 (A4)). The simulations for this case reveal that just by introducing uncertainty in one part of the model (the alternative progression factor) rank reversals take place. A tendency that only becomes stronger as uncertainty is introduced on both levels of the model (alternative and criteria progression factors).

Overall, the results of the simulations show that the perception of the size of the progression factors can have an influence on the ranking of alternatives in situations where the original scores are rather close to each other. This is, however, only the case if the members of the ratifying group have different perceptions of the progression. The simulations for the Rail Baltica case study (Sec. 3.2.1) show that rank reversals are very likely to take place if the alternatives have scores close to each other. This is what would have been expected, and could also have been revealed from more conventional sensitivity analyses which test the sensitivity of criteria weights. However, the overlaps occurring between the two most attractive alternatives for the HH and the Ostlänken cases respectively (Sec. 3.2.2 and 3.2.3) indicate that the issue can occur even if scores are further from each other. The overlaps occurring for the latter case study are rather small, and the calculated probabilities imply that rank reversals only take place if the most extreme values are applied for the progression factors.

This paper has assumed a consensus seeking group that is able to agree on the pair wise comparisons to be made. Instead the approach has focussed on the uncertainty related to the progression factors only. It is, however, very likely that this issue does not stand alone in practical application, and there will often be a need for both addressing the disagreement in the group about the judgments to be made as well as the perception of the scale. Thus in future applications the presented approach could be combined with a technique for addressing the variety in pair wise comparisons between group members such as the ones presented by Van den Honert¹² or Jalao et al.¹³, where individual judgments are modelled using a probability distribution.

Moreover the issue of consistency has not been addressed in this paper, but it is of course not less relevant when dealing with a large group of decision makers. Future applications could also deal with this e.g. by incorporating a method for mitigating judgment contradiction such as suggested by Kou et al.³⁹

5. Conclusion

This paper has proposed a new decision support approach for using stochastic simulation to deal with issues concerning the progression factors in the multiplicative AHP. The paper has focused on the uncertainty associated with the progression factors used when transforming the decision-makers' comparative judgments from the verbal to the

geometric scale in the multiplicative AHP version. Depending on the decision context, the decision-makers' aversion towards risk, etc., it seems most likely that the decision-makers have varying progression factors in their perception of the scale. This issue has not previously been treated in the literature, and thus the paper presents a new approach in form of a decision support technique which is able to model the uncertainty relating to the progression factors in the multiplicative AHP using stochastic simulation. Hereby interval results are obtained instead of point estimates for the scores of the alternatives, and the probabilities of rank reversal between the alternatives caused by the group's different perceptions of the scale are calculated.

The approach reveals that the issue of rank reversals due to different perceptions of the progression factors is most likely to occur in situations where the ratifying group is composed of people with very different world views. In such a situation the mind-sets of the group participants are very different and the scale parameter is likely to vary. This will often be the case if dealing with a decision problem that affects many people, e.g. a transport infrastructure project. In such a case the different stakeholders involved in the decision process could, for example, be officials from the government, local officials from municipalities, environmentalists, representatives from the business community, etc. They will often have very divergent opinions on the project considered and will also have very different qualifications for understanding the scale used for the assessments. This can be overcome by the proposed simulations on the scale parameters, which takes into account of the group's variation in thinking.

The simulations imply that lack of consensus in some instances can lead to a final rank order of the alternatives, which appears non-optimal for a proportion of the group. It is therefore recommended that the group should consider the rank reversal probabilities before selecting an alternative. Hence, individual perceptions of the scale play a significant role, and it can be difficult to determine a "best" alternative from the results of the analysis.

The proposed technique makes it possible for the decision makers to take into account the uncertainty related to the perception of the progression factors. In practise the technique can inform the decision makers if there is a potential risk of overlapping intervals of scores. The issue can then be addressed by a discussion of how large an uncertainty the decision makers are willing to accept, or it can lead to more thorough examinations of the alternatives under consideration e.g. by including additional criteria in the assessment. Hence, the technique can be a useful tool for identifying potential rank reversals between alternatives during the decision making process.

References

1. T.L. Saaty, Scenarios and priorities in transport planning: Application to the Sudan, *Transport Research* **11**(1977) 343-350.
2. F. Zahedi, The analytic hierarchy process – a survey of the method and its applications, *Interfaces* **16**(1986) 96-108.
3. V. Belton, A comparison of analytic hierarchy process and a simple multi-attribute value function, *European Journal of Operational Research* **26**(1986) 7-21.

4. T.J. Stewart, A critical survey on the status of multi-criteria decision making theory and practice, *Omega* **20**(1992) 569-586.
5. P. Leskinen and J. Kangas, Rank reversals in multi-criteria decision analysis with statistical modelling of ratio-scale pairwise comparisons, *Journal of the Operational Research Society* **56**(2005) 855-861.
6. P. Korhonen and H. Topdagi, Performance of the AHP in comparison of gains and losses, *Mathematical and Computer Modelling* **37**(2003) 757-766.
7. J. Barzilai, W. Cook and B. Golany, Consistent weights for judgement matrices of the relative importance for alternatives, *Operations Research Letters* **6:3**(1987) 131-134.
8. J. Barzilai and B. Golany, An axiomatic framework for aggregating weights and weight-ratio matrices, *Proc. Sec. Int. Symp. AHP* (1991) 59-70.
9. J. Barzilai, On the use of the eigenvector in the AHP, *Proc. Int. Conf. MCDM* (1992) 291-300.
10. R.C. van den Honert and F.A. Lootsma, Group preference aggregation in the multiplicative AHP: the model of the group decision process and Pareto optimality, *European Journal of Operational Research* **96**(1996) 363-370.
11. J. Barzilai and F.A. Lootsma, Power relations and group aggregation in the multiplicative AHP and SMART, *Journal of Multi-Criteria Decision Analysis* **6**(1997) 155-165.
12. R.C. van den Honert, Stochastic group preference modelling in the multiplicative AHP: A model of group consensus, *European Journal of Operational Research* **110**(1998) 99-111.
13. E.R. Jalao, T. Wu and D. Shunk, A stochastic AHP decision making methodology for imprecise preferences, *Information Sciences* **270**(2014) 192-203.
14. G. Dede, T. Kamalakis and T. Sphicopoulos, Convergence properties and practical estimation of the probability of rank reversal in pairwise comparisons for multi-criteria decision making problems, *European Journal of Operational Research* **241**(2015) 458-468.
15. H. Eskandari and L. Rabelo, Handling uncertainty in the analytic hierarchy process: a stochastic approach, *International Journal of Information Technology and Decision Making* **6**(2007) 177-189.
16. D. Jato-Espino, J.R. Rodriguez-Hernandez, V.C. Andrés-Valeri and F. Ballester-Muñoz, A fuzzy stochastic multi-criteria model for the selection of urban pervious pavements, *Expert Systems with Applications* **41**(2014) 6807-6817.
17. M.B. Barfod and S. Leleur, Scaling transformations in the REMBRANDT technique: A sensitivity examination of the progression factors, *International Journal of Information Technology & Decision Making* **12**(5)(2013) 1-17.
18. K.B. Salling, Assessment of transport projects: risk analysis and decision support, PhD thesis, Department of Transport, Technical University of Denmark (2008).
19. J.A. Rodger, P. Pankaj and S.P. Gonzalez, Decision making using a fuzzy induced linguistic ordered weighted averaging approach for evaluating risk in a supply chain, *International Journal of Advanced Manufacturing Technology* **70**(2014) 711-723.
20. D.L. Olson, G. Fliedner and K. Currie, Comparison of the REMBRANDT system with analytic hierarchy process, *European Journal of Operational Research* **82**(1995) 522-539.
21. F.A. Lootsma, The REMBRANDT system for multi-criteria decision analysis via pair wise comparisons or direct rating, Report 92-05, Faculteit der Technische Wiskunde en Informatica, Delft University of Technology, Netherlands (1992).
22. R. Ramanathan, Stochastic decision making using multiplicative AHP, *European Journal of Operational Research* **97**(1997) 543-549.
23. Y-M. Wang and T.M.S. Elhag, An approach to avoiding rank reversal in AHP, *Decision Support Systems* **42**(2006) 1474-1480.
24. D.L. Olson, *Decision Aids for Selection Problems* (Springer-Verlag, New York, 1996).

25. S. Zahir, Eliciting ratio preferences for the analytic hierarchy process with visual interfaces: A new mode of preference measurement, *International Journal of Information Technology & Decision Making* **5**(2006) 245-261.
26. F.A. Lootsma, *Multi-criteria decision analysis via ratio and difference judgement* (Kluwer Academic Publishers, Dordrecht, 1999).
27. F.A. Lootsma, Scale sensitivity in the multiplicative AHP and SMART, *Journal of Multi-Criteria Decision Analysis* **2**(1993) 87-110.
28. A.M. Law and W.D. Kelton, *Simulation modelling and analysis* (McGraw-Hill, Singapore, 2000).
29. K.B. Salling and D. Banister, Assessment of large transport infrastructure projects: the CBA-DK model, *Transportation Research Part A* **43**(2009) 800-813.
30. K.B. Salling and S. Leleur, Modelling of transport project uncertainties: feasibility risk assessment and scenario analysis, *European Journal of Transport Infrastructure Research* **12**(2012) 21-38.
31. D. Vose, *Risk Analysis – A Quantitative Guide* (Wiley and Sons Ltd, Chichester, 2008).
32. W.E. Back, W.W. Boles and G.T. Fry, Defining triangular probability distributions from historical cost data, *Journal of Construction Engineering and Management* **126**(2000) 29-37.
33. K.B. Salling and S. Leleur, Assessment of transport infrastructure projects by the use of Monte Carlo Simulation: The CBA-DK model, *Proceedings of the Winter Simulation Conference* (2006) 1537-1544.
34. L.D. Phillips, *Decision Conferencing*, in: W. Edwards, R. F. Miles, D. von Winterfeldt (Eds.), *Advances in Decision Analysis*, Cambridge University Press, New York, 2007, pp. 375-399.
35. I. Ambrasaitė, M.B. Barfod and K.B. Salling, MCDA and Risk Analysis in Transport Infrastructure Appraisals: The Rail Baltica Case, *Procedia Social and Behavioral Sciences* **20**(2011) 944–953.
36. S.T. Petursson, Analysing the sensitivity of the progression factors in the multiplicative analytic hierarchy process, Master Thesis, Department of Transport, Technical University of Denmark (2013).
37. M.B. Barfod and K.B. Salling, A new composite decision support framework for strategic and sustainable transport appraisals, *Transportation Research Part A* **72**(2015) 1-15.
38. M.B. Barfod, A.V. Jensen and S. Leleur, Examination of decision support systems for composite CBA and MCDA assessments of transport infrastructure projects, *Lecture Notes in Economics and Mathematical Systems* **648**(2011) 167-176.
39. G. Kou, D. Ergu and J. Shang, Enhancing data consistency in decision matrix: Adapting Hadamard model to mitigate judgment contradiction, *European Journal of Operational Research* **236**(2014) 261–271