Comparison of Judgment Scales of the Analytical Hierarchy Process - A New Approach

Klaus D. Goepel [a]

[a] BPMSG
Business Performance Management Singapore
2 Bedok Reservoir View #17-02, Singapore 479232
E-mail drklaus@bpmsg.com
2018-05-20

Abstract

The analytic hierarchy process (AHP) remains a popular multi-criteria decision method. One topic under discussion of AHP is the use of different scales to translate judgments into ratios. The author makes a new approach to compare different scale functions and to derive a recommendation for the application of scales. The approach is based on simple analytic functions and takes into consideration the number of criteria of the decision problem. A correction to the so-called balanced scale is proposed, and a new adaptive-balanced scale introduced. Scales are then categorized and compared based on weight boundaries, weight ratio, weight uncertainties, weight dispersion and number of decision criteria. Finally a practical example of a decision hierarchy is presented applying the different scales. The results show that the corrected balanced scale improves weight dispersion and weight uncertainty in comparison to the original AHP scale. The proposed adaptive-balanced scale overcomes the problem of a change of the maximum weight depending on the number of decision criteria.

Keywords

Decision support systems, Multi-criteria decision making, Analytic Hierarchy Process, Judgment scales, Pairwise Comparisons.

1. Introduction

Despite all academic discussions, the analytic hierarchy process (AHP) remains one of the most popular multi-criteria decision making methods (MCDM). Originally proposed by Saaty (1980), over the last decades several modifications and improvement have been proposed. A review of the main developments in AHP can be found from Ishizaka & Labib (2011). One of the topics being under discussion for a long time is the fundamental AHP scale. Saaty and Vargas (2012) describe ratio scales, proportionality and normalized ratio scales as one of the seven pillars of the Analytic Hierarchy Process. The fundamental AHP scale of absolute numbers is derived from the psychophysical law of Weber–Fechner and uses absolute numbers 1, 2, 3 … 9 or its verbal equivalents (Table 1).

Paired comparisons are made by identifying the less dominant of two elements and using it as the unit of measurement. One then determines, how many times more the dominant member of the pair is than this unit. The reciprocal value is used for the comparison of the less dominant element with the more dominant one.

Judgment x	**Verbal equivalent**	**Comment**
1	*Equal importance*	Two activities contribute equally to the objective.
2	Weak or slight	
3	*Moderate importance*	Experience and judgment slightly favor one activity over another.
4	Moderate plus	
5	*Strong importance*	Experience and judgment strongly favor one activity over another.
6	Strong plus	
7	*Very strong or demo-strated importance*	An activity is favored very strongly over another; its dominance demonstrated in practice.
8	*Very, very strong*	
9	*Extreme importance*	The evidence favoring one activity over another is of the highest possible order of affirmation.

Table 1. Fundamental AHP judgment scale with integers 1 to 9 and their verbal equivalents (Saaty, 2008).

Theoretically there is no reason to be restricted to these numbers and verbal gradation, and several other numerical scales have been proposed. They are summarized, based on Ishizaka & Labib (2011), in table 2. There are no guidelines, what scale to use for a specific decision problem, and the choice of the "best" scale is an ongoing discussion. Franeka and Kresta (2014) classified the judgment scales based on consistency and allocation of priorities for a specific example with seven criteria, using the row geometric mean method (RGGM), also known as the logarithmic least-squares method (Crawford and Williams, 1985).

The objective of the current study is to find a set of parameters allowing a further classification of AHP scales, and to analyze and discuss the impact of different scales on the resulting priorities, in order to support the selection of an appropriate scale for AHP projects.

3. Methodology

The author focuses on intangible judgments on the fundamental scale (table 1), translated into ratios $1/M \dots 1 \dots M$. We do not consider actual measurements like distance, area or temperature, where a limited integer scale is not necessary, and AHP can be applied using actual measured ratios.

A first obvious approach is to categorize the scales into the following three categories (table 2):

- *Category 1* – The maximum entry value in the decision matrix is kept at nine: fundamental AHP scale, inverse linear scale, balanced scale and generalized balanced scale
- *Category 2* - The maximum range of entry values in the decision matrix is reduced to lower values than nine: logarithmic scale, root square scale, Koczkodaj scale.
- *Category 3* –The maximum entry values and range of entry values in the decision matrix is extended to values higher than nine: power scale, geometric scale, adaptive and adaptive-balanced scale.

C	No	Name	Short	Scale function c(x)	*M*	Comment
1	1	Fundamental AHP scale	AHP	$c = \mathrm{x}$	9	Saaty (1980)
	2	Inverse linear scale	Inv-lin	$c = \dfrac{9}{10-x}$	9	Ma-Zheng (1991)
	3	Balanced scale	Bal	$c = \dfrac{9+x}{11-\mathrm{x}}$	9	Saalo, Hämäläinen (1997) for [01, 0.9]
	4	Generalized balanced scale*	Bal-n	$c = \dfrac{9+(n-1)x}{9+n-x}$	9	Generalized balanced scale
2	5	Logarithmic Scale	Log	$c = \log_a(x+\mathrm{a}-1)$	3.3	Ishizaka et. al (2010) a=2
	6	Root square scale	Root	$c = \sqrt[2]{\mathrm{x}}$	3	Harker, Vargas (1987)
	7	Koczkodaj scale	Kocz	$c = 1 + \dfrac{x-1}{9-1}$	2	Koczkodaj (2016)
3	8	Power scale	Power	$c = \mathrm{x}^2$	81	Harker, Vargas (1987)
	9	Geometric scale	Geom	$c = a^{x-1}$	256	Lootsma (1989), a=2 Dong et al. (2008)
	10	Adaptive scale*	Adapt	$c = x^{\left(1+\frac{\ln(n-1)}{\ln 9}\right)}$	M^*	$M^* = M(n-1)$
	11	Adaptive-balanced scale*	Adapt-bal	See eq. 9d	M^*	$M^* = M(n-1)$

Table 2. AHP scales investigated in this paper. x is the value on the integer judging scale for pairwise comparisons from 1 to 9, c the ratio used as entry into the decision matrix, M the maximum value of c for $x = 9$. Scales 4, 10 and 11 marked with * are introduced and explained in this paper.

Salo & Hämäläinen (1997) pointed out that the integers from 1 to 9 yield local weights, which are not equally dispersed. They state that for a given set of priority vectors $\boldsymbol{w}_{A\mathrm{HP}}$ the corresponding ratios r can be computed from the relationship

$$r = \frac{w_{AHP}}{1-w_{AHP}} \tag{1a}$$

or

$$w_{\mathrm{AHP}} = \frac{r}{r+1} \tag{1b}$$

In fact, eq. 1a or its inverse eq. 1b are a special case for one pairwise comparison of two criteria. If we take into account the complete n x n decision matrix for n criteria, the resulting weights for a criterion, judged x-times more important than all others, can be calculated as (see annex 1):

$$r = \frac{w_{AHP}}{1-w_{AHP}}(\mathrm{n}-1) \tag{2a}$$

$$w_{\mathrm{AHP}} = \frac{r}{r+n-1} \tag{2b}$$

Preprint of an article submitted for consideration in International Journal of Information Technology and Decision Making

Eq. 2b simplifies to eq. 1b for n=2.

We will use eq. 2b to calculate the weights $\boldsymbol{w}_{\mathbf{AHP}}$ for different number of criteria n and different scales by substituting r with the scale functions $\boldsymbol{c} = \boldsymbol{f}(\boldsymbol{x})$ of table 2. In our numerical examples the number of criteria n is varied from n =2 to n = 9. This range follows the recommendation to keep the maximum number of criteria in the range of the magic number seven plus or minus two (Saaty, T., Ozdemir, M. S., 2003) and covers most of the practical applications.

We then investigate all scales looking at the following common parameters:

1. *Weight bound and weight ratio*: What is the *maximum weight* for a judgment that one criterion is "9 - extreme more important" than all others, and how compare the total *ratios* of calculated weights for different scales?
2. *Weight uncertainty*: How much depend the weights on *small variations of the judgement*?
3. *Weight dispersion*: How are the weights distributed over the judgment range?

Parameter 1 gives us the range of possible weights as a function of the number of criteria. Parameter 2 reflects the rounding when using an integer judgment scale; the resolution is 1, therefore variations of $\Delta = x \pm 0.5$ would reflect the same judgment, but result in weight variations $\pm\Delta w$. Parameter 3 shows, how evenly dispersed (balanced) the weights are over the weight range.

4. Data Analysis and Discussion

We first will have a closer look at the balanced scale (table 2, no. 3) proposed by Saalo and Hämäläinen (1997). The scale was designed in a way that local weights are evenly dispersed over the weight range [0.1, 0.9]. Based on eq. 1a it is computed as

$$c = \frac{w_{\text{bal}}}{1 - w_{\text{bal}}} \quad (3a)$$

with

$$w_{\text{bal}} = 0.45 + 0.05\,x \quad (3b)$$

For $x = 1 \ldots 9$ weights are equally distributed from 50% to 90%.

The balanced scale can be written as $$c = \frac{9 + x}{11 - \text{x}} \quad (3c)$$

c (resp. $1/c$) are the entry values in the decision matrix, and x the pairwise comparison judgment on the 1 to 9 scale. As shown before, eq. 1a is the special case for one selected pairwise comparison of two criteria. We now use eq. 2a to formulate the more general case of the balanced scale for n criteria and a judgment x with x from 1 to M, resulting in evenly dispersed weights:

$$c = \frac{w_{\text{bal}}}{1 - w_{\text{bal}}} (\text{n} - 1) \quad (4a)$$

With evenly dispersed weights

$$w_{\text{bal}}(x) = w_{\text{eq}} + \left[\frac{w_{\max} - w_{\text{eq}}}{M - 1}\right](x - 1) \quad (4b)$$

using

$$w_{\text{eq}} = \frac{1}{n} \quad (4c)$$

Preprint of an article submitted for consideration in International Journal of Information Technology and Decision Making

and
$$w_{\max} = \frac{M}{n+M-1} \tag{4d}$$
$$w_{\min} = \frac{1}{n+M-1} \tag{4e}$$

we get the generalized balanced scale as

$$w_{\text{bal}} = \frac{1}{n} + \left[\frac{\frac{M}{n+M-1} - \frac{1}{n}}{M-1}\right](x-1) \tag{5a}$$

Setting M = 9
$$w_{\text{bal}} = \frac{1}{n} + \left[\frac{9}{8(n+8)} - \frac{1}{8n}\right](x-1) \tag{5b}$$

the generalized balanced scale can be written as

$$c = \frac{9+(n-1)x}{9+n-x} \tag{5c}$$

We see that eq. 5c with n=2 represents the classical balanced scale as given in eq. 3c. We call eq. 5c the generalized balanced scale or balanced-n (bal-n) scale.

In order to compare AHP weights as a function of the judgments x with the number of criteria n as parameter, we use eq. 3c in eq. 2b to reflect the actual weights of the classical balanced scale for more than 2 criteria (n > 2).

Balanced scale
$$w_{\text{AHP}} = \frac{x+9}{(2-n)x+11n-2} \tag{6a}$$

AHP fundamental scale
$$w_{\text{AHP}} = \frac{x}{x+n-1} \tag{6b}$$

Generalized balanced scale
$$w_{\text{AHP}} = \frac{9+(n-1)x}{n(n+8)} \tag{6c}$$

Fig. 1 visualizes the three functions for n = 7 criteria. It can be seen that a single judgement "5 – strong more important" yields to a weight of 45% on the AHP scale, 28% on the balanced scale and 37% on the generalized balanced scale. Equations 6a, b and c show that, compared to the generalized balanced scale, criteria are underweighted using the classical balanced scale and over weighted using the fundamental AHP scale. Only for n = 2 the classical balanced scale is identical with the generalized balanced scale and yields evenly distributed weights.

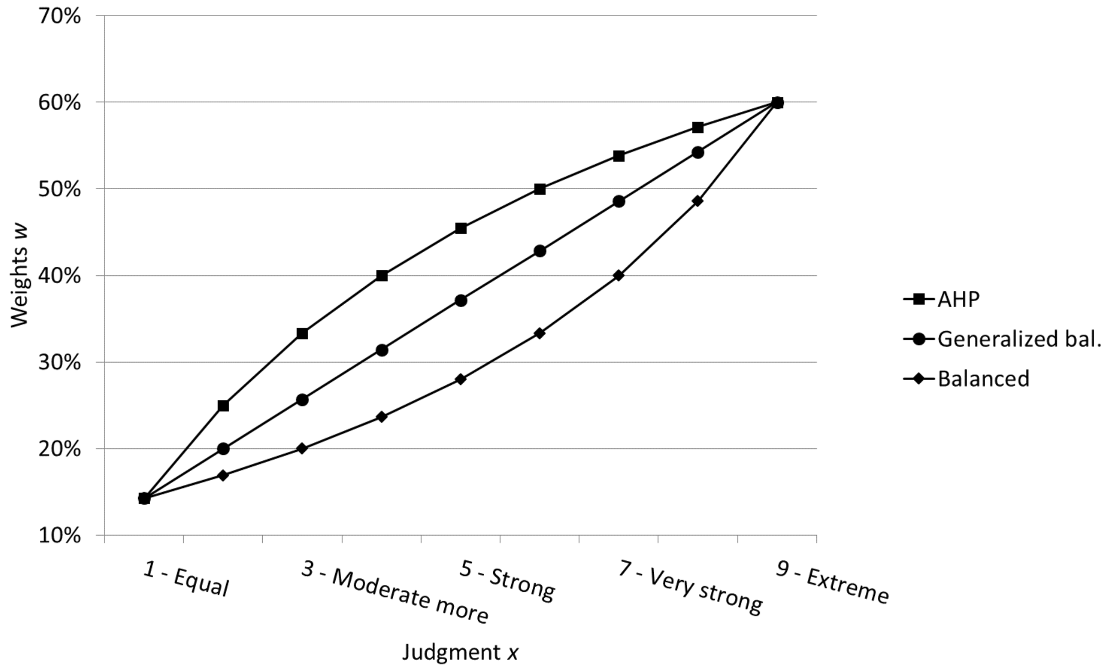

Figure 1. Visualization of eq. 6a, 6b and 6c: weights as function of judgment x for the fundamental AHP scale, the balanced scale and the generalized balanced scale for n = 7 decision criteria.

4.2 AHP adaptive scales

All published AHP scales under study in table 2 are functions of the judgement x and are not depending on the number of criteria. We have shown that the number of criteria n has an impact on the result (eq. 2b). We now can design a scale, where we keep the weight of the extreme most important criterion (eq. 4d) at a constant value over the number n of criteria, and where we will have a constant weight range w_{max}/w_{min} for all n. We call this an adaptive scale.

We calculate M^* to keep the maximum weight w_{max} at 90% for all possible n:

$$\frac{M}{n+M-1} = 0.9 = w_{\max} \tag{7a}$$

$$M^* = M(n-1) \tag{7b}$$

We choose the scale function c as a function of x to have the form $c = x^y$ in order to keep it linear with the logarithm of the stimulus x. With $c_{max} = M^* = M(n\text{-}1)$, $c_{min} = x = 1$ and $M = 9$ we get:

$$y = 1 + \frac{\ln(n-1)}{\ln(9)} \tag{8a}$$

and as a result the adaptive scale

$$c = x^{1+\frac{\ln(n-1)}{\ln 9}} \tag{8b}$$

For n = 2 eq. 8b represents the original AHP scale, for n = 10 it represents the power scale. For all n the maximum possible weight is 90%.

The same concept can be applied to the generalized balanced scale (eq. 5c) using w_{max} = 0.9, and we get the adaptive-balanced scale with:

$$w_{\text{bal}} = \frac{1}{n} + \frac{0.9 - \frac{1}{n}}{8}(x-1) \tag{9a}$$

$$w_{\text{bal}} = \frac{(9n-10)(x-1)+80}{80n} \tag{9b}$$

$$c = \frac{w_{\text{bal}}(x)}{1-w_{\text{bal}}(x)}(n-1) \tag{9c}$$

Adaptive balanced scale

$$c = \frac{(9n-10)(x-1)+80}{(9n-10)x-89n+90}(n-1) \tag{9d}$$

This scale function keeps the maximum weight at 90%, independent from the number of criteria; at the same time the weights are equally distributed over the range [0.1, 0.9] as for the generalized balanced scale.

4.3 Weight boundaries and weight ratios

A shortcoming of any scale with a finite upper bound M is that the upper bound restricts the range of local priority vectors. The weight boundary w_{max} is given by eq. 4d. Scales in category 3 (table 2) extend the upper bound M from nine to higher values, scales in category 2 reduce the upper bound M from nine to lower values. For example, the highest priority for two criteria is 90% on the fundamental AHP scale, 75% on the root square scale and 98.8% on the power scale. The maximum weight w_{max} decreases with increasing number of criteria (eq. 4d). The resulting weights for the same judgement 9 - extreme more important can vary from 90% (AHP scale, $n = 2$ criteria) to 52% (AHP scale, $n = 9$ criteria).

When judging a specific criterion as "extremely more important" (9 on the fundamental AHP scale) than all other criteria, a decision maker would expect the resulting weight for this particular criterion to come out "significantly higher" than all other weights. Let us quantify the case: if the weight for the extreme most important criterion has a value of less than 50%, it is in fact no longer the extreme most important criterion, as all other weights together already exceed the weight of the extreme most important criterion.

We calculate the weight ratio WR of different AHP scales as the ratio of the maximum weight for one "extreme most important" criterion over the sum off all other weights.

$$WR = \frac{w_{\max}}{(n-1)w_{\min}} \tag{10a}$$

With eq. 4d and e we get

$$WR = \frac{M}{n-1} \tag{10b}$$

The fundamental AHP scale has a weight ratio of nine, or one order of magnitude for two criteria. Scales in category 3 ($M > 9$) expand the weight ratio and improve the discrimination of weights, scales of category 2 ($M < 9$) compress the weight ratio, and weights come closer together.

If we set the threshold at a weight ratio of one, *i.e.* the most important criterion gets the same weight as the rest of criteria, we get

$$\frac{M}{n-1} = 1 \text{ or simply } n = M + 1 \tag{11}$$

The fundamental AHP 1 to 9 scale and the other scales in category 1 cross the threshold at n = 10 criteria. Scales of category 2 cross the threshold at three or four criteria, and scales of category 3 cross the threshold at ten or more criteria. The adaptive scales in category 3 have a constant weight ratio of 9, independent from the number of criteria *n*. Interestingly, Saaty's fundamental 1 to 9 scale seems to represent a kind of compromise: for two criteria the extreme most important criterion receives a weight of 90%; most decision makers would probably accept it as a fair representation of their judgment. The threshold of the fundamental AHP scale is reached at the number of ten criteria, which is close to the recommendation, to keep the maximum number of criteria in the range of the *Magical Number Seven plus or Minus Two* (Saaty & Ozdemir, 2003).

Based on the weight boundary and weight ratio we can now compare the scales, up to which number n_{max} of criteria they are applicable under the condition that the extreme most important criterion gets a weight of at least 50% (table 3).

Scale	AHP	Inv-lin.	Bal-n	Log	Root	Kocz	Power	Geom
n_{max}	10	10	10	4	4	3	82	257

Table 3. Maximum number of criteria n_{max} for different AHP scales based on a weight threshold of 50% for a single extreme most important criterion.

It can be seen that for practical applications of medium complexity (decision problems with more than three or four criteria) the introduced threshold eliminates the use of the scales in category 2 for more than four criteria.

4.4 Weight uncertainty

We will now look at small variations of judgements by Δx and their impact on the resulting weights Δw_{AHP}. This will give us an idea about uncertainties of the resulting weights. Using an integer judgment scale $x = 1 \dots 9$, a judgment of, for instance, $x = 3$ covers an interval from 2.51 to 3.49, if we interpret x as a rational number. We determine the weight variation $\Delta w_{\text{AHP}}(x)$ based on eq. 2b with $r = c(x)$ and $r = c\ (x \pm \Delta x)$ using the scale functions $c(x)$ from table 2 for the different scales.

$$\Delta w_{\text{AHP}} = w_{AHP}[c(x + \Delta x)] - w_{AHP}[c(x)] \qquad (12a)$$

The derivative of eq. 2b gives us

$$\Delta w_{\text{AHP}} = \frac{n-1}{[c(x)+n-1]^2} \cdot \frac{dc(x)}{dx} \cdot \Delta x \qquad (12b)$$

For each value on the integer judgment scale we take the derivative eq. 12b at $x \pm \Delta x/2$ with $\Delta x = 0.5$, corresponding to half the resolution of the integer judgment scale. In the case of the fundamental AHP scale with $c = 1$ the derivative is

$$\Delta w_{\text{AHP}} = \frac{n-1}{(x+n-1)^2} \cdot \Delta x \qquad (13a)$$

The weight uncertainty Δw_{AHP} has its maximum at $x = 1$. For n=3 criteria we get

$$\Delta w_{\text{AHP}} = \frac{2}{(1+0.25+3-1)^2} \cdot 0.5 = 9.5\% \qquad (13b)$$

Judging all three criteria as equal important using the fundamental AHP scale results in equal local weights of 33.3% with an uncertainty of 9.5%.

Analyzing the weight uncertainties for all scales with eq. 12b (see table 4), we find that the fundamental AHP scale, logarithmic scale, root square scale and power scale show the highest uncertainties for *x* = 1 (equal importance), decreasing with increasing *x* and increasing *n*.

Scale	Derivative $\Delta w_{\text{AHP}}/\Delta x$
Fundamental AHP scale	$\frac{n-1}{(x+n-1)^2}$
Inverse linear scale	$\frac{9(n-1)}{\left((n-1)x-10n+1\right)^2}$
Balanced scale	$\frac{20(n-1)}{((n-2)x-11n+2)^2}$
Generalized balanced scale	$\frac{n-1}{n(n+8)}$
Logarithmic scale	$\frac{\ln(2)(n-1)}{(x+1)(\ln(x+1)+\ln(2)(n-1))^2}$
Root square scale	$\frac{n-1}{(\sqrt{x}+n-1)^2}\cdot\frac{1}{2\sqrt{x}}$
Koczkodaj scale	$\frac{n-1}{\left[\frac{x-1}{8}+n\right]^2}\cdot\frac{1}{8}$
Power scale	$\frac{2x(n-1)}{(x^2+n-1)^2}$
Geometric scale	$\frac{\ln(2)(n-1)}{(2^x+2n-2)^2}\cdot 2^{x+1}$
Adaptive scale	$\frac{(\ln(n-1)+\ln(9))(n-1)x^{\ln(n-1)/\ln(9)}}{\ln(9)\left(x^{1+\frac{\ln(n-1)}{\ln(9)}}+n-1\right)^2}$
Adaptive-balanced scale	$\left(\frac{9}{80}-\frac{1}{8n}\right)$

Table 4. Derivatives of weights as function of judgments *x* for different scales used to calculate weight uncertainty and weight dispersion.

The power scale has the highest, the root square scale the lowest weight uncertainties. The inverse linear scale shows the highest uncertainty for *x*=9 (extreme importance). The geometric scale shows a maximum at

$$x=\frac{\ln(2n-2)}{\ln(2)} \tag{14a}$$

The maximum moves from *x* = 1 (equal importance) for *n* = 2 criteria to *x* = 4 (moderate plus importance) for *n* = 9 criteria. The adaptive scale lies in between the fundamental AHP scale (*n*=2) and the power scale (*n*=10). For the generalized balanced scale and the adaptive balanced scale uncertainties do not depend on the judgment; they are constant over the judgment range and can be calculated as

Generalized balanced scale $$\Delta w_{\text{AHP}} = \frac{n-1}{n(n+8)} \cdot \Delta x \qquad (14b)$$

Adaptive balanced scale $$\Delta w_{\text{AHP}} = \left(\frac{9}{80} - \frac{1}{8n}\right) \cdot \Delta x \qquad (14c)$$

Fig. 2 shows the behavior for *n* = 7 parameter.

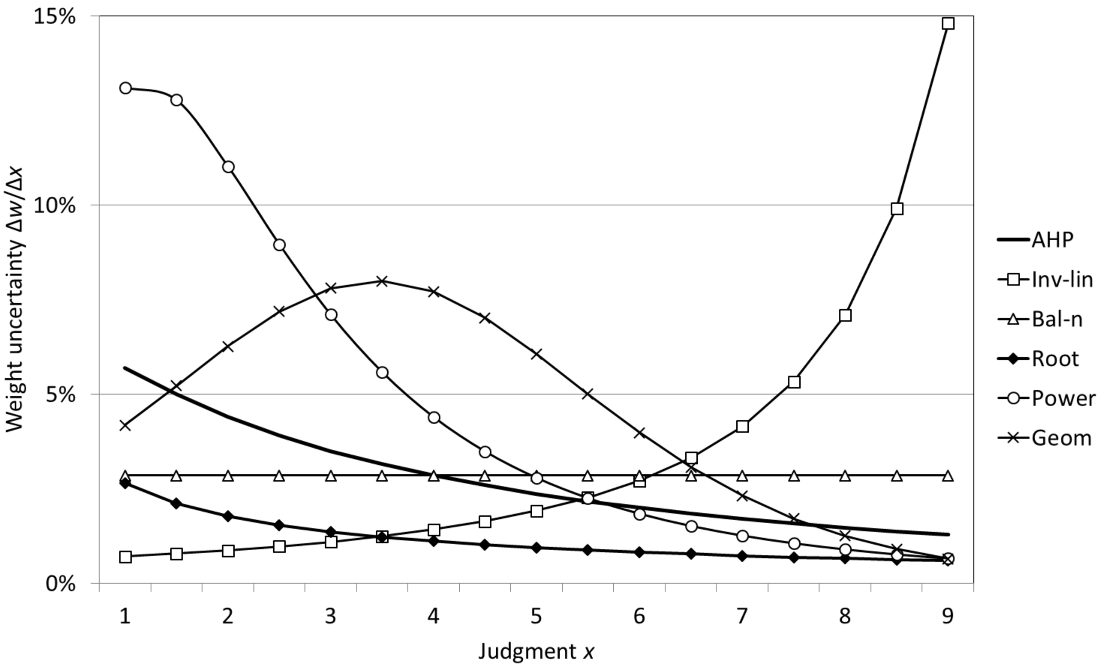

Figure 2. Weight uncertainty as a function of judgments (*x* = 1 to 9) for different AHP scales and *n* = 7 criteria.

As we can see from the above, it is important for all scales to identify the weight uncertainties, as they are not negligible, and they could have an impact on the results for some decision problems.

4.5 Weight dispersion

The integers *x* = 1 to 9 yield local weights, which are unevenly dispersed. For example, a judgment change from *x* = 1 to 2 yields to a weight change of 17%, whereas a judgment change from *x* = 8 to 9 results in a weight change of only 1.1%; a factor of 15-times less. There is a lack of sensitivity, when comparing elements close to each other. This was a reason for Saalo and Hämäläinen (1997) to introduce the balanced scale. Ishizaka *et al.* (2010) proposed a logarithmic scale, which is smoother for high values. Their work does not support the Saaty scale, but they prefer a more balanced scale, like the balanced scale or the inverse linear scale.

As a measure of weight dispersion *WD* for different AHP scales we calculate the standard deviation of the differences of weights *w* for each transition on the 1 to 9 judgment scale

$$WD = \sqrt{\sum(\Delta w_i - \overline{\Delta w})^2/7} \qquad (15)$$

with $$\Delta w_i = w_{x=i} - w_{x=i+1}$$

for i = 1 ... 8.

The weight differences Δw_i are calculated using the derivatives from table 4. Evenly or more uniform distributed priorities will give a lower standard deviation than unevenly distributed weights. As we can see from table 4, Δw does not depend on the judgment x for the generalized balanced and adaptive balanced scale, the weight dispersion WD is zero. Scales in category 2 (Koczkodaj, logarithmic and root square) have a lower, scales in category 3 (power, adaptive) a higher weight dispersion than the fundamental AHP scale (figure 3). For $n = 2$ the adaptive scale has the same weight dispersion as the fundamental AHP scale and for $n = 10$ the same weight dispersion as the power scale. For all scales the weight dispersion decreases with increasing number of criteria.

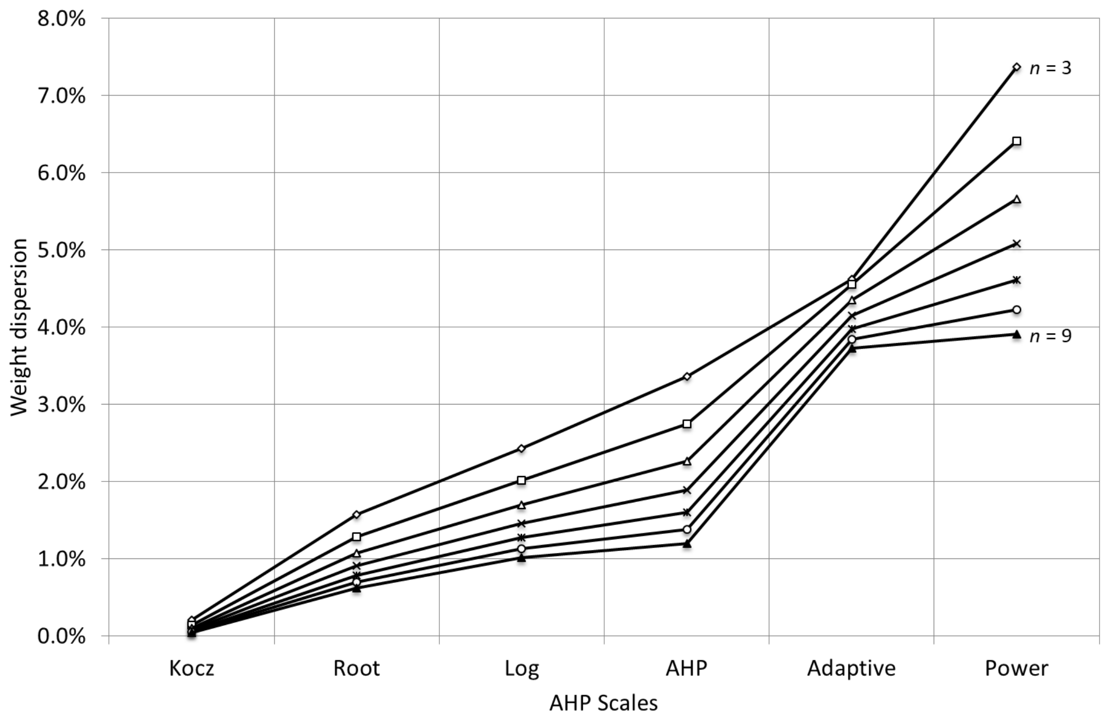

Figure 3. Weight dispersion for different AHP scales. Parameter is the number of decision criteria from n=3 to n=9.

The geometric and the inverse-linear scales show a different behaviour. For the geometric scale WD is practically independent from the number of criteria, and with $WD \approx 2.6\%$ the weight dispersion lies between the logarithmic and the fundamental AHP scale. The weight dispersion of the inverse-linear scale increases with increasing number of criteria and is higher than the weight dispersion of the fundamental AHP scale for $n > 4$ criteria (see fig. 4).

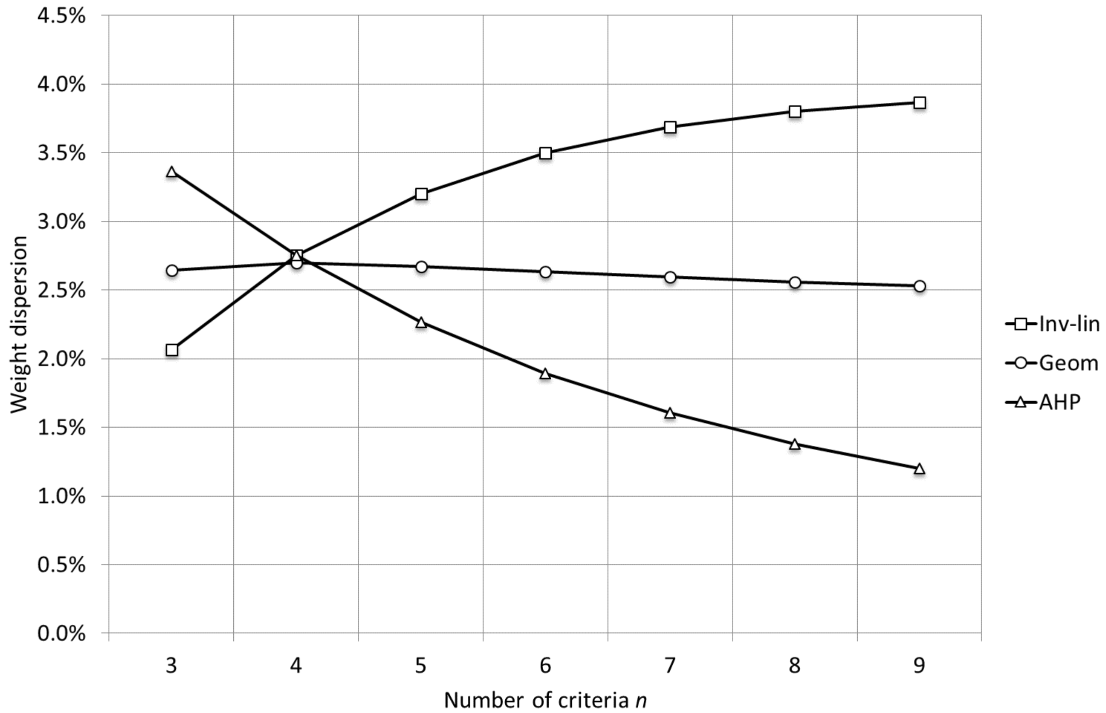

Figure 4. Weight dispersion for the geometric and inverse-linear scale compared to the fundamental AHP scale as function of number of criteria *n*.

4.6 Comparison of AHP scales

We can now compare and discuss all scales based on the parameters described in the previous paragraphs:

Max. number of criteria n_{max} (eq. 11) $n_{\text{max}} = M + 1$

Weight boundary (eq. 4d) $w_{\text{max}} = \frac{M}{n+M-1}$

Weight ratio (eq. 10b) $WR = \frac{M}{n-1}$

Weight Uncertainty (eq. 12b) $\Delta w_{\text{AHP}} = \frac{n-1}{[c(x)+n-1]^2} \cdot \frac{dc(x)}{dx} \cdot \Delta x$

Weight Dispersion (eq. 15) $WD = \sqrt{\sum(\Delta w_i - \overline{\Delta w})^2/7}$

In the comparison table (table 5) we select n =2 and n = n_{max} criteria and show the maximum of the weight uncertainty over the whole judgment range x = 1 to 9.

For *category 1* scales weight boundary, weight range and the max. number of criteria are the same. Differences can be seen in the max. *weight uncertainty* and *weight dispersion*. By concept, the generalized balanced scale has no weight dispersion, weights are equally distributed over the judgment range. The original AHP scale has a lower weight dispersion and slightly lower uncertainty than the inverse-linear scale. Based on weight uncertainty and weight dispersion the generalized balanced scale is preferable compared to the original AHP scale.

<table>
<tr><th rowspan="2">Cat</th><th rowspan="2">Scale</th><th rowspan="2">n_{max}</th><th colspan="2">Weight boundary %</th><th>Weight Ratio</th><th colspan="2">Max. Weight Uncertainty %</th><th colspan="2">Weight Dispersion %</th></tr>
<tr><th>n=2</th><th>$n= n_{max}$</th><th>n=2</th><th>n=2</th><th>$n= n_{max}$</th><th>n=2</th><th>$n= n_{max}$</th></tr>
<tr><td rowspan="4">1</td><td>Fundamental AHP</td><td rowspan="4">10</td><td rowspan="4">90</td><td rowspan="4">50</td><td rowspan="4">9</td><td>9.9</td><td>4.3</td><td>3.1</td><td>1.1</td></tr>
<tr><td>Inverse-linear</td><td>4.7</td><td>16</td><td>1.1</td><td>3.9</td></tr>
<tr><td>Balanced</td><td>2.5</td><td>7.8</td><td>0</td><td>2.0</td></tr>
<tr><td>Generalized balanced</td><td colspan="2">2.5</td><td colspan="2">0</td></tr>
<tr><td rowspan="3">2</td><td>Logarithmic</td><td rowspan="2">4</td><td>77</td><td>53</td><td>3.3</td><td>6.8</td><td>5.5</td><td>2.1</td><td>2.0</td></tr>
<tr><td>Root square</td><td>75</td><td>50</td><td>3</td><td>5.0</td><td>4.0</td><td>1.4</td><td>1.3</td></tr>
<tr><td>Koczkodaj</td><td>3</td><td>67</td><td>50</td><td>2</td><td>1.5</td><td>1.4</td><td>0.3</td><td>0.2</td></tr>
<tr><td rowspan="4">3</td><td>Power</td><td rowspan="4">≥10</td><td>99</td><td>90</td><td>81</td><td>19</td><td>11</td><td>6.2</td><td>3.6</td></tr>
<tr><td>Geometric</td><td>99.6</td><td>96.6</td><td>256</td><td>6</td><td>8</td><td>2.3</td><td>2.5</td></tr>
<tr><td>Adaptive</td><td colspan="2">90</td><td>9</td><td>10</td><td>11</td><td>3.1</td><td>3.6</td></tr>
<tr><td>Adaptive-bal</td><td colspan="2">90</td><td>9</td><td>2.5</td><td>4.9</td><td colspan="2">0</td></tr>
</table>

Table 5. AHP scale comparison

The critical point for all *category 2* scales is the compression of the weight ratio. It yields to a less significant discrimination of weights, and based on the threshold for the maximum number of criteria, they should not be applied for problems with more than three or four criteria.

Category 3 scales expand the weight range and make the discrimination of priorities more significant. The geometric scale is preferable compared to the power scale, as it has a lower weight uncertainty and also a lower weight dispersion. The newly proposed adaptive-balanced scale combines a higher weight range with low uncertainty and equally distributed weights.

Comparing the scales across all categories, generalized balanced and adaptive-balanced scale show the best values. A further advantage is that their weight uncertainty is constant over the whole judgment range 1 to 9, and the uncertainty does not exceed 5% for up to ten criteria.

4.7. Implementation and practical example

We will now show an example of a realistic project to demonstrate the findings of this paper using a free web based AHP online system (AHP-OS), which allows to switch between different scales (Goepel, 2014). The software also estimates weight uncertainties using randomized variations of all judgments by ±0.5 on the judgment scale. The example is taken from Saaty (1990), “Buying a house”, because all necessary input data are given in his paper. The decision matrix has eight criteria. The calculated weights are shown in table 6.

No	Criterion	w_{AHP} %	Rank
1	Size of house	17.3	3
2	Transportation	5.4	5
3	Neighbhood	18.8	2
4	Age of house	1.8	8
5	Yard space	3.1	7
6	Modern facilities	3.6	6
7	Gen. condition	16.7	4
8	Financing	33.3	1

Table 6. Priorities and ranking of criteria for the example taken from Saaty (1990)

The criterion *Financing* (8) has the highest weight of 33%, three criteria (1, 3, 7) have a weight of approx. 18% and the remaining four criteria (2, 4, 5, 6) have weights from 2% to 5%. Figure 5 shows the change of these weights, when applying the different AHP scales discussed in this paper. The scales are sorted by ascending weight ratio of the resulting priorities.

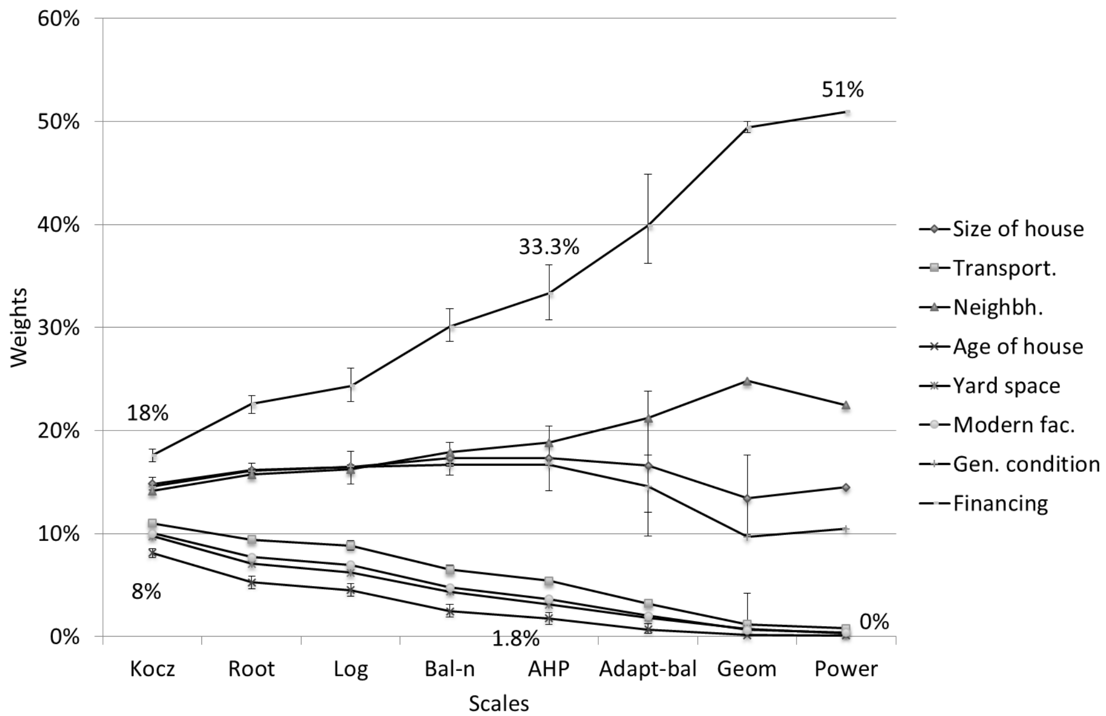

Figure 5. Changes of weights for the example with eight criteria as a function of different AHP scales. The error bars indicate the weight uncertainty based on a randomized variation of judgment values by ±0.5 on the judgment scale.

Scales of category 2 *compress* the weight range, scales of category 3 *expand* the weight range. The weight range for the Koczkodaij scale (highest compression) is 10%, *i.e.* all calculated priorities lie between 8% and 18%. For the power and geometric scale the range expands to 50%.

The weight of the criterion with the highest weight (8, Financing) changes from 18% on the Koczkodaij scale to 51% on the power scale. Weights in the mid-range change less under different scales, weights in the low range increase for category 2 scales, and decrease for category 3 scales.

Comparing our results with the example given by Franeka & Kresta (2014) with 7 criteria, we also see a high variance of allocation of priority values for the power and geometric scale, a moderate variance for the fundamental AHP, generalized balanced and balanced scale, and a low variance for all scales in category 2.

Saaty (1990) also shows in his example the evaluation of three alternatives, house A, B, C. Table 7 compares the results using different scales. Weights of criteria and alternatives are both evaluated using the same scale. Fundamental AHP, adaptive-balanced and generalized balanced scale results are close with no change of the ranking of alternatives; root square and geometric scale show a change in the ranking.

Scale	**Alternative A**		**Alternative B**		**Alternative C**	
	w %	Rank	w %	Rank	w %	Rank
AHP	39.6	1	34.1	2	26.3	3
Adapt-bal	39.7	1	34.1	2	26.2	3
Bal-n	40.8	1	30.6	2	28.6	3
Root	38.4	1	30.1	**3**	31.5	**2**
Geom	40.5	2	43.6	**1**	15.9	**3**

Table 7. Alternative evaluation for example from Saaty (1990) under different scales.

Consistency Ratio CR

Franeka & Kresta (2014) calculated new random indexes for the different scales in order to calculate the consistency ratio. In our example we used the standard calculation as proposed by Saaty (1980). For the above example *CR* is with 17% already higher than the recommended threshold of 10% for the fundamental AHP scale. Applying the different scales, it increases for category 3 scales, and decreases for category 2 scales. Although the author did not investigate the impact of the scales on the consistency ratio *CR* in further detail, it seems that scales of category 2 lower *CR* and scales of category 3 increase *CR*. In category 1 the inverse-linear and generalized balanced scale lower *CR* compared to the fundamental AHP scale.

5. Limitations

In order to get an analytical expression for the priority vector as a function of judgments, we have investigated the particular case that one criterion is x-times more important than all others using the row geometric mean method. We only considered consistent inputs and did not study the effect of small perturbations with nearly consistent matrices for different scales. The example presented in 4.7, as well as the work of Franeka & Kresta (2014) confirm the findings in practical projects with nearly consistent matrices, but more case studies are necessary. In addition, the impact of different scales on the consistency ratio *CR* requires further investigations.

6. Conclusion

The discussion about the *right scale* for the analytic hierarchy process is ongoing for many years. With this paper the author has shown that it is possible without complex mathematics or computer simulations (*e.g.* Dong et al, 2008) to derive some fundamental relations for the evaluation of different AHP scales. Based on the case that a decision maker judges a single criterion x-times more important than n-1 others, we can derive a simple analytical relation (eq. 2b) to calculate the priority vector using the row geometric mean method.

Due to the limitation of the scale to a maximum judgment value (usually nine on the fundamental AHP scale), we can also calculate the weight boundaries (eq. 4d, e) and the maximum weight ratio (eq. 10b).

In a first step it was then shown that the balanced scale has to be generalized in order to take into account the number of criteria and to yield equally distributed priorities across the judgment range. A modification of the classical balanced scale was presented and the generalized balanced scale (eq. 6c) introduced.

AHP scales were categorized in three categories, depending on the maximum entry value to the decision matrix. For a final comparison of scales *weight boundaries*, *weight ratio*, *weight uncertainties* and *weight dispersion* over the judgment range were used. To overcome the limitations of the maximum weight depending on the number of criteria, an adaptive-balanced scale (eq. 9d) was proposed and included in the comparison. In addition to the theoretical calculations a typical decision example was evaluated using the different scales of this study.

Based on the analysis shown and parameters discussed the main findings can be summarized as follows.

1. The so-called balanced scale has to be generalized and has to take into account the number of criteria in order to be applied for more than two criteria. Otherwise local priorities will not be balanced and will be underweighted compared to the generalized balanced scale and the fundamental AHP scale.

2. Scales reducing the entry ratio into the decision matrix to values lower than nine (category 2) *compress* the calculated weights, making weight discrimination more difficult. Based on a threshold of 50% for one single most preferred criterion their application for decision problems with more than three or four criteria is not recommended.

3. Scales extending the entry ratio into the decision matrix (category 3) *expand* the calculated weights, making weight discrimination easier. At the same time they show higher weight dispersion and weight uncertainties increase.

4. The fundamental AHP scale seems to present a kind of compromise with respect to the maximum number of criteria, weight dispersion and weight uncertainty. For all category 1 scales only the generalized balanced scale improves weight dispersion and weight uncertainty in comparison to the original AHP scale. Practical projects indicate an improvement of the consistency ratio *CR* for the generalized balanced scale.

5. The proposed adaptive-balanced scale overcomes the problem of a change of the maximum weight depending on the number of criteria. It keeps the weight ratio at nine for any number of criteria and results in evenly distributed weights across the judgment range, but shows higher values of weight uncertainties than the generalized balanced scale.

7. Acknowledgments

The development of the free AHP online software system AHP-OS (Goepel, 2014) started in 2014. Over the years we received many questions and feedback from its users. The different scales, discussed in this study, were implemented just recently based on their feedback. It gave us the motivation to take a closer look at the scale problem and publish this paper.

References

Crawford G, & C, W. (1985). *A Note on the Analysis of Subjective Judgement Matrices*, Journal of Mathematical Psychology, 29, 387-405.

Dong, Y., Xu, Y., Li, H., Dai, M. (2008). *A comparative study of the numerical scales and the prioritization methods in AHP*, European Journal of Operational Research 186, p 229–242.

Franeka, J. , Kresta, A. (2014). *Judgment Scales and Consistency Measure in AHP*, Procedia Economics and Finance 12:164–173.

Goepel, K. D. (2014). *AHP Online System (AHP-OS), Rational decision making made easy*, http://bpmsg.com/academic/ahp, accessed 17.6.2017.

Harker, P., & Vargas, L. (1987). *The Theory of Ratio Scale Estimation: Saaty's Analytic Hierarchy Process*. Management Science, 33, 1383-1403.

Ishizaka, A., Balkenborg, D., & Kaplan, T. (2010). *Influence of aggregation and measurement scale on ranking a compromise alternative in AHP*. Journal of the Operational Research Society, 62, 700-710.

Ishizaka, A., Labib A. (2011). *Review of the main developments in the analytic hierarchy process*, Expert systems with applications, 38(11), 14336-14345.

Koczkodaj W.W. (2016) *Pairwise Comparisons Rating Scale Paradox*. In: Nguyen N.T., Kowalczyk R. (eds) Transactions on Computational Collective Intelligence XXII. Lecture Notes in Computer Science, vol 9655, 1-9. Springer, Berlin, Heidelberg.

Lootsma, F. (1989). *Conflict Resolution via Pairwise Comparison of Concessions*. European Journal of Operational Research, 40, 109-116.

Ma, D., & Zheng, X. (1991). 9/9-9/1 *Scale Method of AHP*. In 2nd Int. Symposium on AHP (Vol. 1, pp. 197-202). Pittsburgh.

Salo, A.,Hämäläinen, R. (1997). *On the measurement of preferences in the analytic hierarchy process*, Journal of multi-critria decision analysis,Vol. 6, 309 – 319.

Saaty, T. L. (1980). *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*, ISBN 0-07-054371-2, McGraw-Hill.

Saaty, T. L. (1990). *How to make a decision: The Analytic Hierarchy Process*, European Journal of Operational Research 48, 9-26.

Saaty, T.L. (2003). *Decision-making with the AHP: Why is the principal eigenvector necessary*, European Journal of Operational Research 145, 85–91.

Saaty, T.L., Ozdemir, M. S. (2003). *Why the magic number seven plus or minus two*, Mathematical and computer modelling, 38 233-244.

Saaty, T. L. (2008). *Decision making with the analytic hierarchy process*, Int. J. Services Sciences, Vol. 1, No. 1, 2008, p 83 - 98

Saaty, T. L., Vargas, L. G. (2012). *The seven pillars of the analytic hierarchy process*, Models, Methods, Concepts & Applications of the Analytic Hierarchy Process, International series in operations research & management science 175, p 23 – 40, Springer ISBN 978-1-4614-3596-9 .

Annex 1: AHP weights as a function of judgments

Let *DM* be a *n* x *n* decision matrix, where the first criterion is *x*-times more important than all others. Then the first matrix element is "1", and the rest of the first matrix row is filled with (n-1)-times *x*. The first matrix column is filled with (n-1)-times 1/x.

$$DM = \begin{pmatrix} \mathbf{1} & x & x \\ 1/x & \mathbf{1} & 1 \\ 1/x & 1 & \mathbf{1} \end{pmatrix} \tag{a1}$$

To calculate the priorities, we use the Row Geometric Mean Method (RGGM), as the decision matrix is consistent and the result will be the same as for the right eigen vector.

$$\text{RGGM} \rightarrow \begin{pmatrix} (x^{n-1})^{1/n} \\ \left(\frac{1}{x}\right)^{1/n} \\ \left(\frac{1}{x}\right)^{1/n} \end{pmatrix} \tag{a2}$$

The resulting weights (priorities) for the first criterion is the normalized geometric mean of the first row.

$$w_{\text{AHP}} = \frac{(x^{n-1})^{\frac{1}{n}}}{(x^{n-1})^{\frac{1}{n}} + (n-1)(x^{-1})^{\frac{1}{n}}} \tag{a3}$$

With some rearrangement

$$w_{\text{AHP}} = \frac{1}{1 + \frac{(n-1)(x^{-1})^{\frac{1}{n}}}{(x^{n-1})^{\frac{1}{n}}}} = \frac{1}{1 + \frac{(n-1)x^{-\frac{1}{n}}}{x \cdot x^{-\frac{1}{n}}}} = \frac{1}{1 + \frac{(n-1)}{x}} \tag{a4}$$

we get the simple relation

$$w_{\text{AHP}} = \frac{x}{x+n-1} \tag{a5}$$

□