
Supplier evaluation process by pairwise comparisons

Arkadiusz Kawa · Waldemar W.
Koczkodaj

Abstract In this study, we propose to assess suppliers by using consistency-driven pairwise comparisons for tangible and intangible criteria. The tangible criteria are simpler to compare (e.g. the price of a service is lower than that of another service with identical characteristics). Intangible criteria are more difficult to assess. The proposed model combines assessments of both types of criteria.

The main contribution of this paper is the presentation of an extension framework for the selection of suppliers in a procurement process. The final weights are computed from relative pairwise comparisons. For the needs of the article, surveys were conducted among Polish managers dealing with cooperation with suppliers in their enterprises. The Polish practice and restricted bidding are discussed, too.

Keywords Suppliers selection · assessment model · supply chain · pairwise comparison · inconsistency analysis

1 Introduction

The success of an enterprise does not only depend on all the co-operating subjects [38]. Efficiency of a given subject as a whole is not its basic source of competitive advantage. It is the efficiency of various types of activities that the enterprise undertakes when delivering its product to the market [6]. These

A. Kawa
Department of Logistics and Transport
Poznan University of Economics, Poznan, Poland
E-mail: arkadiusz.kawa@ue.poznan.pl

W.W. Koczkodaj
Department of Mathematics and Computer Science
Laurentian University, Sudbury, Ontario, Canada P3E 2C6
E-mail: wkoczkodaj@cs.laurentian.ca

actions create a supply chain. The main objective of the supply chain is to provide the maximum value to customers at low costs and high speed [3,19].

An important issue of a supply chain management is the procurement [?,40]. The cost of products and services acquired from external suppliers is significant for most manufacturing firms [15]. On average, manufacturers purchases of goods and services amount to 55% of revenues and it is in contrast to labor costs of 6% and overhead expenses of 3% of revenues [33]. For high-technology firms, purchased material and services represent up to 80% of total product costs [5,35].

The objective of the procurement process must be the harmonization of internal processes of buyers and suppliers in order to avoid a waste of resources within the supply chain. It may be achieved when a great amount of emphasis is put on establishing and maintaining good relations with suppliers. Optimization in the area of delivery may bring a chance of sizable savings. It is noteworthy that the scale of a company's activity may result in greater losses caused by incorrect purchasing processes.

Reliable suppliers enable manufacturers to reduce inventory level and improve product quality, which is the main reason why concerns about appropriate suppliers is increasing. One of the prime responsibilities of the purchasing function is the evaluation and selection of suppliers. Some researchers even indicate that evaluation and selection of suppliers are critical due to their contribution to supply chain performance [1].

Enterprises build a base of suppliers with whom they cooperate more or less closely for many years. When searching for solutions related to the choice of a new supplier or evaluation of an existing one, people responsible for such decisions take into account a range of criteria. These criteria are very often worked out on the basis of the company's experience.

Carefully selected, competitive suppliers can go a long way in minimizing adverse impacts and, in fact, in enhancing positive impacts on the quality of the output of an organization. The importance of an adequate framework for the selection of suppliers for an organization has been stressed in the literature [24,27].

The supplier selection process requires predetermination of several issues, including the number of suppliers with which the organization wishes to contract for a given product or service [13]. It is difficult to determine how many suppliers should work for a given organization. Some authors argue that there is now a trend to reduce the number of suppliers to a manageable level [29,30], even to a single source in extreme cases.

No clear recommendation regarding the number of suppliers is confirmed by the results of the research. Every third company does not follow any general rule. This is understandable. On the one hand, a large number of suppliers ensures lower prices for customers, offers greater safety and decreases the risk of stopping production. On the other hand, it raises the operating costs of such cooperation (maintenance costs of information systems, controlling, sourcing, negotiation, setting conditions for the cooperation, audits, etc.) [8,9].

In the process of partner selection, several methods are used. Among them, categorical method, weighted-point method, vendor performance matrix approach, vendor profile analysis (VPA), multi-criteria decision aid (MCDA), multiple objective programming (MOP) such as goal programming, data envelopment analysis (DEA) and multi-attribute utility theory (MAUT). The recent studies about the use of fuzzy theories and their development in the supplier selection problem are particularly interesting [16,39,17]. In [26] the authors proposed a new tool for supplier selection based on a combination of the grey system and the uncertainty theory neither of which requires any probability distribution or fuzzy membership function.

The structure of the paper is as follows: the characteristic of bidding process is discussed in Section 2. Then, Section 3 describes supplier evaluation problems and Section 4 presents pairwise comparisons method and review of the other applications and literature on the pairwise comparisons. Section 5 provides example of a tendering model proposed by the authors. Next, Section 6 contains consistency analysis of tendering model. Section 7 proposes an example of suppliers assessment. The conclusions of research are outlined in Section 8.

2 Bidding process

The previously mentioned methods have some limitations because the supplier selection process is mostly based on intuition. There is no theoretical base or consistent method of predicting the best bid. It is not uncommon for the evaluation panel to arrive at a deadlock when a part of the panel favors one solution, because of certain criteria, while the other part insists on another solution since it scores better on different criteria. The decision making process nearly always involves some kind of constituency in modern democratic societies. We have various boards of governors or directors, committees, task groups, city councils, panels of experts, and individuals, each with a specific agenda. Heated discussions and various ways of dispute and argumentation often take place to arrive at certain decisions.

Most constituencies have worked out precise and practical policies for running meetings in an orderly and effective manner. What we lack, however, is a device for drawing solid consistent conclusions and all too often the loudest individual wins! Unfortunately, loudness does not necessarily go along with wisdom. Casual thinking is not efficient in predicting complex outcomes.

The main goal of tendering is the selection of the most suitable supplier from the company point of view. Through a bidding we try to achieve:

- the setting of common input constraints for potential suppliers,
- the selection of the best supplier based on tangible and intangible but constant (during the entire bidding process) criteria, which allow us to compare the proposed offers,
- a minimization of the influence of informal interests on selection of an offer thanks to the application of a strict selection process.

One of the most instrumental conditions of a fair bidding is the necessity of a precise scoring system of all criteria and their preferences in the tendering documents to be used during the selection process by the selection panel.

A selection panel is obligated to use objective and measurable evaluation criteria during the selection process. The goal of such an action is to have more comparability and fewer intangible or arbitrary decisions. Application of such quantification in the request for proposals (RFPs) to specific bidding components, together with the bidding amount, helps in the selection of the best offer. A recommended practice is the assignment of weights (e.g., factors which are positive points) to each aspect of an offer with respect to the criteria, listed in the bidding documents. As a result, the best offer is the one with the highest total score.

In restricted bidding (as opposed to open bidding), only invited parties may submit a bid. In the first stage of a two stage bidding process suppliers submit preliminary bids without the total amount. However, the selection panel may request precise specification of some parameters. Successful bidders from the first stage are invited to submit a complete bid for the second stage. Two-stage bidding may be considered a special case of an open bidding. The main differences are:

- the possibility of submitting two tendering offers by bidders,
- the use of two selections instead of one; the first selection to pre-qualify tenders, the second to find the true winner,
- the possibility (and sometimes necessity) of negotiations with bidders,
- the possibility of using the negotiation results, to change or restrain crucial constraints of the request for proposals, before they are distributed again to bidders selected during the first stage of bidding.

The two-stage bidding process is recommended in situations:

- where it is difficult to predict certain parameters such as: technical, quality, service, or construction tasks,
- where necessity demands negotiating with suppliers or constructors because of the specific character of the supplies and/or construction services,
- when the tendering is related to research, assessment, or any other specialized service.

3 Is there an easy way to evaluate potential suppliers?

Evaluation problems are not new. They have deep roots in our modern perception of measurement systems. One of the major achievements of our technical civilization is standardization. It is fair to say that without standards, rapid technological progress would have not taken place. Standardization should not only be perceived as a “nuts and bolts” or physical concept (without which new products and spare parts would be far more expensive), but also as an intellectual concept. In the case of “nuts and bolts”, it is a concept of the standard thread pitch. Today, we use standards to such a degree, that we often

forget about other alternatives. The problem is that we often assume that we have a measure for nearly everything! No one questions the practicality of the measure of length (meter or foot), or weight (kg or pound), as they are used in everyday life. There are, however, situations where it is more useful and natural to use pairwise comparisons. Consider this example of the reliability of two suppliers A and B. The reliability of supplier A is evaluated as 4 on the scale of 0 to 5 and B is 2 on the same scale. There is no standardized unit of reliability hence it is easier to say that A is two times more reliable than B and it is what we call “pairwise comparison”. Interestingly the use of a standardized measure unit such as a meter stick is also a pairwise comparison. The statement “The length of A is 4m” is an abbreviation of “By a pairwise comparison of A to one meter we have a factor of 4”.

Using pairwise comparisons (see, for example, [28,36,7]) is natural and it does not need to be imprecise. We have become so accustomed to having standards, that sometimes we find it difficult to imagine a situation where no standard measures exist. The truth is that there are many such situations.

Measurement of the environment or environmental pollution are good examples of situations where a standard yardstick seems to be missing. For example, it would be hard to use a cubic meter of environment as a standard measure since in one cubic meter of environment there could be millions of ants but only a fraction of an elephant. How could one decide if a fraction of an elephant is less significant than a colony of ants?

4 Pairwise comparisons

Casual thinking is partial, fragmentary, and is not an effective way to measure intangibles. In the decision making process, many factors must be considered simultaneously and with about the same degree of importance. As such, an approach with more finesse is necessary to obtain a clear and unambiguous conclusion. It has been shown by numerous examples that the pairwise comparisons method can be used to draw final conclusions in a comparatively easy and elegant way. The brilliance of the pairwise comparisons could be reduced to a rule of common sense: *consider two objects at a time if you are unable to handle more than that*. Lull noticed it in 13th century.

The practical and theoretical virtue of the pairwise comparisons methodology is its simplicity. The goal of pairwise comparisons is to establish the relative preference of two criteria in situations in which it is impractical (or sometimes meaningless) to provide the absolute estimations of the criteria. To this end, an expert (or a team of experts) provides relative comparison coefficients $a_{ij} > 0$, which are meant to be a substitute for the quotients s_i/s_j of the unknown (or even undefined) absolute values of criteria $s_i, s_j > 0$. The quotients s_i/s_j are also sometimes called *relative weights* in the literature.

Code	Definition of intensity or importance	Application
1	Equal importance	Two criteria equally contribute to the objective or lack of knowledge to compare them
2	Essential or strong importance	Experience and judgments favor one criterion over another
3	Absolute importance	The highest affirmation degree of favoring one criterion over another
1.4, etc.	Intermediate judgments	When compromise is needed

Table 1 Comparison scale

$$\mathbf{A} = \begin{vmatrix} 1 & a_{12} & \cdots & a_{1n} \\ \frac{1}{a_{12}} & 1 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{a_{1n}} & \frac{1}{a_{2n}} & \cdots & 1 \end{vmatrix}$$

where a_{ij} expresses an expert's relative preference of criteria s_i , over s_j .

Coefficients a_{ij} are expected to satisfy some natural restrictions (e.g., $a_{ii} = 1$, $a_{ij} \cdot a_{ji} = 1$). For the sake of our exposition we define the pairwise comparisons $n \times n$ matrices simply as square matrices $A = (a_{ij})$ such that $a_{ij} > 0$ for every $i, j = 1, \dots, n$.

A pairwise comparisons matrix A is called *reciprocal* if $a_{ij} = \frac{1}{a_{ji}}$ for every $i, j = 1, \dots, n$ (then automatically $a_{ii} = 1$ for every $i = 1, \dots, n$). Even a stronger condition seems natural. A pairwise comparisons matrix A is called *consistent* if $a_{ik} = a_{ij} \cdot a_{jk}$ for every $i, j, k = 1, \dots, n$. While every consistent matrix is reciprocal, the inverse in general fails. Consistent matrices correspond to the ideal situation, in which there are exact values s_1, \dots, s_n for criteria. The quotients $a_{ij} = s_i/s_j$ form a consistent matrix. Conversely, the starting point of the pairwise comparisons inference theory, which states that for every $n \times n$ consistent matrix $A = (a_{ij})$, there exists positive real numbers s_1, \dots, s_n such that $a_{ij} = s_i/s_j$ for every $i, j = 1, \dots, n$. Such vector $s = (s_1, \dots, s_n)$ is unique up to a multiplicative constant.

How can we establish fair weights? Is there a theory to help us? The weighting classification needs to be done on a fair basis for every criteria, which ought to have its share in contributing to the overall judgment. A fair solution is to compare all criteria in pairs using, for example, a small scale from 1 to 3 (as mathematically reasoned, on the basis of the Fulop's constant, in [12]), presented in Table 1. The solution accuracy of not-so-inconsistent matrices depends on the inconsistency (see [14] for details). Smaller scale contributes to decreased inconsistency. One may note that the consistency-driven approach is, in brief, the next step forward in the development of pairwise comparisons. By concentrating on the consistency as the means to improve the precision of judgments it picks up where Analytic Hierarchy Process arrived to. AHP inconsistency index detected the existence of the inconsistency by the eigenvalue-based heuristic and its rather questionable threshold of 10%

randomly generated matrices. In [21], a mathematical reasoning (based on two counterexamples) was provided to reject the eigenvalue-based inconsistency.

It is not our goal to present the consistency-driven approach here, but only its application to evaluation of proposals. The theoretical foundations of this method are based on [22, 11, 34], while [23] presents convincing statistical evidence that the pairwise comparisons are contributing to the improvement of accuracy. The pairwise comparisons methodology was introduced by Thurstone, in 1927 (see [34]). Its further extension by the consistency-driven analysis (see the next section as well as [22]) can be employed as a powerful theoretical framework for the evaluation of tenders.

One of the notable pairwise comparisons past applications of national importance is the evaluation of site proposals for nuclear power plants in Holland (rejected by Dutch government, for details see [28]). The Monte Carlo study presented in [14] discounts a claim of superiority of any particular method of solving the pairwise comparisons matrix. In fact, the accuracy of solutions strongly depend on the consistency of judgments making the consistency-driven approach suitable for applications.

A new definition of inconsistency introduced in [22] is instrumental for the inconsistency analysis. This definition allows us to locate the most inconsistent judgments and is instrumental for extending the hierarchical model of pairwise comparisons by the inconsistency analysis.

Issues related to public bids are documented in [2]. Some of the legal aspects of bids are outlined in [10]. Multiattributes (multicriteria) evaluation and its relation to the analytic hierarchy process is presented by some authors. None of them, however, presents a comprehensive approach focused only on the bidding process and the evaluation of overall of criteria. Our approach includes all the aspects.

5 Example of tendering model

When the company already has a specific set of suppliers, with whom it intends to cooperate, it should choose the best of the many suppliers, characterized by various features. The vendor selection process must, therefore, be properly designed. Tools that help with the decision on the selection of a partner, and thus allow to assess bids from multiple suppliers and compare them and point to the best ones, are necessary.

Supplier evaluation is based on a set of both tangible and intangible criteria. The former are relatively easy to compare (e.g., the price of a product is lower than that of another product with identical characteristics). The latter is more difficult to survey and require detailed considerations.

A practical model of a tendering process needs to be as flexible as possible (see also conclusions). Presenting any model for such process is risky, since some readers may conclude at this point that it is not suitable for him/her. One may always discount any model as irrelevant. However, leaving a reader without any practical application of the presented framework is unacceptable.

Tender evaluation criteria	
Quality of product or service (<i>qua</i>)	Reliability and stability of the product or service (<i>reliability</i>)
	Quality standards (<i>q-standards</i>)
	Quality guarantees (<i>q-guaranty</i>)
	Quality control system (<i>q-system</i>)
Flexibility and adaptability (<i>flex</i>)	Ease of change introduction (e.g., of resources, processes) (<i>change</i>)
	Readiness to execute orders according to clients terms (<i>ready</i>)
	Ability to deal with a changing economic environment and crisis situations (<i>a-deal</i>)
	Ability to produce personalized products or improve existing ones according to customers' requirements (<i>a-personal</i>)
	Ability to cooperate and coordinate in a completely new product (<i>a-cooperate</i>)
	Reaction time to unexpected demand changes (<i>reaction</i>)
Organizational potential (<i>pot</i>)	Production capacity (<i>product</i>)
	Logistics opportunities (<i>logist</i>)
	Experience in cooperation (<i>cooper</i>)
	Management and organizational skills (<i>skills</i>)
Financial standing and payment conditions (<i>fin</i>)	Loan possibilities (<i>loan</i>)
	Cash conversion cycle (<i>cash</i>)
	Financial stability (<i>stable</i>)
	Capital rotation (<i>capital</i>)
Experience (<i>exp</i>)	Number of customers (<i>customers</i>)
	Number of completed transactions (<i>trans</i>)
	Recommendation other customers (<i>recommend</i>)
	Supported industries (<i>industry</i>)

Table 2 Criteria taken into consideration in the selection and evaluation process of suppliers

Therefore a compromised solution is proposed and Table 2 contains a set of selected criteria most frequently used in tenders. These criteria are based on the indicators in the SCOR documentation [4,32], articles, studies, reports about selections criteria, analyses of companies and interviews with experts. Thus, our model is a mixture of a lot of approaches to supplier selection. We chose some criteria of supplier selection which were verified by experts in supply chains. The proposed list is not exhaustive and does not pretend to be complete but still fairly representative. It is worthwhile to note that the authority issuing a tender can select arbitrary criteria for each type of bid. They can also scale (or weight) particular criteria, depending on the kind (or extent) of work, the required potential of the contractor, or the necessary level of technology.

The criteria have been divided into five main groups:

- quality of product or service (*qua*),
- flexibility and adaptability (*flex*),
- organizational potential (*pot*),
- financial standing and payment conditions (*fin*),

- experience (*exp*).

The above mentioned criteria together form a Supplier Quality Index (SQI). The two criteria: price and delivery time are not mentioned here deliberately. These are the tangible factors, which can be compared with each other easily [?], but they are usually threshold factors. We could, for example, get a very good deal from a supplier who could deliver, what we need now, in ten years. Similarly with the price. For this reason, our model includes factors which are usually labeled *other factors*. For this reason, our model should be regarded as an extension to the classical model of supplier selection since the computed factor may be used to top of what we are forced to do it estimating *other factors*. It is all well reflected by the *satisficing rule* introduced by Herbert A. Simon in 1956 (see [31]). Moreover, Simon is a winner of Nobel Prize in Economics (1978). Besides, the tangible factors are usually the criteria of the greatest weight in the evaluation process. SQI is, therefore, an important addition to the analysis of suppliers.

Obviously, when choosing suppliers, all types of criteria can be considered. Often, they form close relationships. For example, price is inextricably linked to the quality of the product or service. However, immeasurable factors such as payment terms and supplier flexibility have an indirect effect on the price of the product or service.

The criteria presented in Table 2 may not always meet the needs of the company. Therefore, they should be regarded as a starting point to determine its own set of criteria, that a specific company will use. This means that this list is not exhaustive and, if required, can be extended or shortened. It shall be noted, however, that the adopted criteria should be known to all the suppliers from the enterprise network being evaluated.

For the needs of the article, surveys were conducted among managers dealing with co-operation with suppliers in their enterprises. The study was divided into two stages. First, the entrepreneurs were interviewed in order to make a list of the most important criteria for selecting suppliers in a supply chain. On this basis, a questionnaire was prepared for the previously mentioned survey. The second stage of the study (with the use of the questionnaire) was aimed at establishing the importance of the individual criteria (by comparing them in pairs) for selecting suppliers in a supply chain.

The survey was conducted in September and October of 2013. A representative sample of 49 entrepreneurs from Poland participated. The respondents included mainly directors, logistics specialists, supply chain managers and merchants. It is worth to note that according to this formula: $Sample\ size = 0.25 \times (certainty\ factor / acceptable\ error)^2$ in [37] showing a table with *certainty factor*, a sample size of 49 observations is sufficient to achieve approximately 85% certainty usually desired for a pilot study.

Every person had to answer 10 questions and, thus, determine 10 ratios: $a_{12}, a_{13}, a_{14}, a_{15}, a_{23}, a_{24}, a_{25}, a_{34}, a_{35}, a_{45}$, which allows the formation of the partial PC matrix A in the form:

$$\mathbf{A} = \begin{vmatrix} 1 & a_{12} & a_{13} & a_{14} & a_{15} \\ \frac{1}{a_{12}} & 1 & a_{23} & a_{24} & a_{25} \\ \frac{1}{a_{13}} & \frac{1}{a_{23}} & 1 & a_{34} & a_{35} \\ \frac{1}{a_{14}} & \frac{1}{a_{24}} & \frac{1}{a_{34}} & 1 & a_{45} \\ \frac{1}{a_{15}} & \frac{1}{a_{25}} & \frac{1}{a_{35}} & \frac{1}{a_{45}} & 1 \end{vmatrix}$$

	qua	flex	pot	fin	exp
qua	1	1.5	1.2	1.4	1.3
flex	0.7	1	0.8	1.1	0.9
pot	0.9	1.2	1	1.4	1.2
fin	0.7	0.9	0.7	1	1.1
exp	0.8	1.1	0.9	0.9	1

Table 3 Relative judgements for the group of criteria

Table 3 demonstrates a matrix with relative comparisons. A scale of 1 to 3 (and its inverse 1/3 to 1) is used. In the represented case, the highest importance has been assigned to *qua* because the quality of product and service (e.g., reliability and stability of the product or service) is usually of great importance to customers. When compared to the flexibility and adaptability (*flex*), *qua* has been assessed to be 1.5 (a compromised evaluation between 1 and 2 from Table 1). Flexibility and adaptability is a factor related to the ability to deal with a changing environment and to the reaction time to unexpected demand changes. It is assumed that organizational potential (*pot*) is less important for the customers (comparison of *qua* to *pot* is 1.2). This allows one to assess such evaluation criteria as production capacity, logistics opportunities, experience in cooperation, management, and organizational skills. The assessment of the importance of the quality factors against the financial standing and payment conditions (*fin*) (e.g., cash conversion cycle, financial stability) is set to 1.4 (when compared to *qua*). The *qua* factor is only 1.3 times more important than the experience (*exp*), which is related to the number of customers, recommendation of other customers, etc. *Flex* against *pot* is assessed to 0.8 which means that *pot* is 1.2 times more important than *flex* (see Figure 1). The relationships between *flex* and *fin*, as well as *flex* and *exp* are quite similar and equal 1.1 and 0.9, respectively. *Pot* against *fin* is set to 1.4 and 1.2 when compared to *exp*. The last comparison is *fin* to *exp* and it equals only 1.1, which means *fin* is almost equally important as *exp*. It is worth to note that the maximum of the relative judgments in our research equals only 1.5. This means that there are no significant differences between the presented criteria. No relative judgment is absolute, or even strongly important in comparison to the others.

Bold face 1's on the main diagonal are arbitrary values due to the fact that they represent a relative ratio of a criterium against itself. Values below the main diagonal do not need to be entered by the user. They are reciprocal to the corresponding values in the upper triangle (for details see matrix **A**).

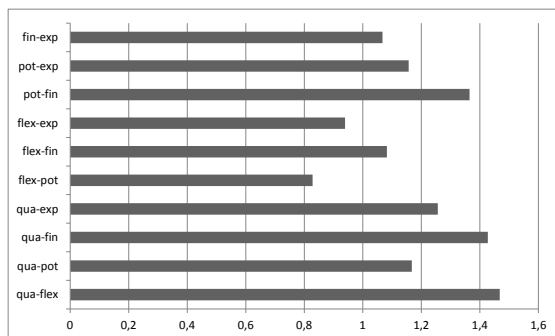


Fig. 1 Relative judgments for the group of criteria on a bar chart

Figure 1 presents data from Table 3 on a bar chart.

6 Consistency analysis of tendering model and final weights for evaluated criteria

One of the challenges posed to the pairwise comparisons method is the lack of consistency of the pairwise comparisons matrices which arise in practice (see [25]). Given an $n \times n$ PC matrix A which is not consistent, the theory attempts to provide a consistent $n \times n$ matrix B which differs from matrix A “as little as possible”. Let $s = (s_1, \dots, s_n)$ be the eigenvector of A corresponding to σ , the largest eigenvalue of A . A statistical experiment (see [14]) shows, however, that the accuracy of weights does not strongly depend on the method. In particular, the geometric means method (see [18]) produced similar results with high accuracy to the ten million cases. There is, however, a strong relationship between accuracy and consistency. This is the main focus of the consistency-driven approach.

In practice, inconsistent judgments are difficult to avoid, when at least three factors are independently compared against each other [21]. In this study, consistency was successfully achieved. For example, let us look closely at the ratios of the first three criteria in Table 3: *qua* (for short A), *flex* (denoted by B), and *pot* (referred as C). The assessment of A against B is 1.5, B against C is assessed as 0.8. The ration of A to C is 1.2.






Let us try to illustrate the inconsistency analysis. From $\frac{A}{B} = 1.5$ and $\frac{B}{C} = 0.9$, we can infer that $\frac{A}{C} = 1.5 \cdot 0.9 = 1.35$, which is different from our original assessments (see Table 3 where $\frac{A}{C} = 1.3$). In fact we do not know which assessment was incorrect. In particular (a frequent case in practice) each original assessment might have been (and usually is) only slightly inaccurate.

x	y	z	cf
1.5	1.3	0.9	0.1
1.2	1.3	1.2	0.1
1.4	1.3	1.1	0.2
0.8	0.9	1.2	0.0
1.1	0.9	1.1	0.2
1.5	1.2	0.8	0.0
0.8	1.1	1.4	0.0
1.5	1.4	1.1	0.1
0.8	0.9	1.2	0.0
1.4	1.2	1.1	0.2

Table 4 Consistency analysis

The consistency factor (cf) is the minimum of $|1 - \frac{1.3}{1.5 \cdot 0.9}|$ and $|1 - \frac{1.5 \cdot 0.9}{1.3}|$ which is 0.04. Since cf is very low, those judgments must not be reconsidered before any further calculations (e.g., of the final weights) can take place. For details related to consistency analysis see [22].

Table 4 shows the consistency analysis for all triads (A, B, C) from the matrix with relative comparisons (see Table 3), where $x = \frac{A}{B}$, $y = \frac{A}{C}$, $z = \frac{B}{C}$. The maximum value of cf is only 0.2, which means that the respondents' answers are consistent enough based on the former application as well as theory provided in [22]. Basically, the heuristic "by one off" and the "satisficing rule" (mentioned in [31]) are used. Simply, the comparisons do not need to be 100% exact, but they should be good enough and "by one off" from the ideally consistent minimal cycle of three comparisons (needed to compare three criteria which is the minimum size of a cycle - two elements do not create a cycle).

Criterion name	Weight	R-tot	Bar graph representation
1. Product or service quality	25%	25%	
2. Financial standing	22%	47%	
3. Organizational potential	18%	65%	
4. Flexibility and adaptability	18%	83%	
5. Experience	17%	100%	

Where R-tot means *running total* of the weight

Figure 2: The final weights for the evaluated criteria

By increasing or decreasing PC matrix entries (the most inconsistent triads are easily located by the simple search procedure), we develop a very good orientation quickly.

All the above computations, including the final weights, are done by the Concluder software. It is freely distributed by SouceForge as Concluder with a tutorial posted on YouTube as jConcluder. It is a flexible and powerful

Evaluation criterion	Weight	Sup. A		Sup. B		Sup. C		Sup. D	
	w_i	x_1	$w_1 * x_1$	x_2	$w_2 * x_2$	x_3	$w_3 * x_3$	x_4	$w_4 * x_4$
Quality of product or service	0.25	3	0.75	4	1	5	1.25	4	1
Organizational potential	0.18	2	0.36	4	0.72	5	0.9	5	0.9
Financial standing	0.22	3	0.66	4	0.88	3	0.66	3	0.66
Experience	0.17	4	0.68	4	0.68	4	0.68	5	0.85
Flexibility and adaptability	0.18	5	0.9	5	0.9	4	0.72	5	0.9
Total	1.00	-	3.35	-	4.18	-	4.21	-	4.31

Table 5 Sample assignment of numerical values to criteria for supplier evaluation

evaluation tool for complex systems with the expectations of having more applications in the future as more decisions are made under growing financial stress.

It is not important to address all mathematical aspects of getting the final weights but the eigenvector method (see [22] for details) can be used to obtain results illustrated in Figure 2. As we can see, the product quality and service factor has the highest weight (25%). The second most important criterion are the financial standing and payment conditions (22%). The other three factors are quite similar and they equal to 17-18%.

One common concern needs to be addressed: *Can we do it?* in short, *Yes, you can.* Using a more precise consistency-driven approach may look complicated at first glance. It may be particularly visible when it comes to making the comparative judgments. How can we start? Is it not just easier to assign points to the list? It is even advisable to start with assigning some points using the so called *by eye* common sense method. Having that done, we can easily construct the pairwise comparisons matrix \mathbf{A} (see Section 6) by simple division of points for the corresponding criteria. In fact, one may even give up at this point, however, after a careful examination of the matrix, we often may be unpleasantly surprised by our own judgments. We may, for example, discover that certain ratios are surprisingly small, while others are out of common sense limits, since comparing two at a time is easier than *by eye* estimation.

7 Example of suppliers assessment

The next stage of the supplier selection is their assessment. Having established the criteria, knowing their weights, one can proceed to the analysis of suppliers in the given circumstances. If these criteria and suppliers are abundant, it may be helpful to apply a multi-criterion analysis using mathematical methods, such as optimization at the so-called multiplicity of attributes. It involves calculation of the synthetic evaluation with the use of the weighted mean, in the way presented below. With a list of individual ratings criteria: x_1, x_2, \dots, x_n with weights: w_1, w_2, \dots, w_n , where $w_n > 0$, the weighted mean can be calculated: $\bar{x} = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i}$.

The weighted mean is calculated for individual suppliers. Finally, those suppliers for which the mean value is highest, are taken into account.

For the needs of the article the authors took into account a large enterprise which produces clothes and is located in Poland. The aim of the company was to implement a supplier evaluation process that would allow to choose the best fabric for the production of clothes from among the suppliers (A, B, C, D). Table 5 presents the criteria from Table 2 supplemented with weights (from Figure 2), assessments and synthetic assessments calculated (assessment multiplied by weight). The grading scale ranges from 1 to 5 (1 being the worst rating and 5 the best). In this case, the best final grade was awarded to supplier D (score 4.31), which, despite lower quality (lower x_i value) than that of supplier C, received better evaluations for most of the other criteria (higher x_i value) from the recipient. If supplier D improved its financial standing and payment conditions, and in particular the loan possibilities, cash conversion cycle, and capital rotation, the result would be even better and it would reinforce its position among other suppliers of the analyzed company. Supplier A got the worst evaluation (score 3.35). This supplier has the worst product or service quality and organizational potential of all.

The authors are aware that the scheme presented in this section may not seem very complicated. However, the procedures offered by the multi-criterion programming should be conceptually simple enough for a person who does not have preparation in mathematics and operations research to use them without difficulty. It is very important, since the evaluation of many criteria (e.g., flexibility and adaptability) cannot take place without the participation of employees, even though most of the steps in this procedure are automated.

A recent application of the presented method for assessment at a national level in Poland is in [20].

8 Conclusions

This study presents a theoretical framework for building a logical model for evaluation of suppliers. It also proposes how the final weights may be obtained from relative pairwise comparisons. However, the framework is not a ready cookbook for constructing such models. Models can be constructed in many different ways. A good model should, however, contain both tangible and intangible criteria. Identification of major criteria is one of the essential components of constructing models. Each model strongly depends on the specifics of the given case and its business environment. The presented approach allows the buyers to create their own list of criteria without any limitations.

The price and delivery time have not been included in our approach. They are threshold factors and reserved for the final decision by the evaluation panel. In other words, the model provides the assessments of tenders as a vector of weights. These weights are applied to evaluation of each individual proposal. The final product of the evaluation generates a list of tenders with the overall score (received by a vector product of weights and corresponding evaluations

for each criteria), plus the price and the delivery time. Frequently, it is an evaluation panel's decision whether or not to award the contract to the best tender according to the overall score.

There are easy cases when the winner's proposed price and delivery time is similar to the next, but in other cases it may be substantially different. Simply, the financial constraints may prohibit consideration of some solutions (e.g., the delivery time may be unacceptable). The consistency-driven pairwise comparisons method can be used by everyone on a personal computer thanks to aforementioned Concluder software.

More research is needed to shortlist all intangible criteria which are used to evaluate suppliers. Currently, it is done by the intuition but it needs to be investigated. The proposed model can process any number of such criteria.

The authors intend to carry out studies on a larger sample of managers who will represent not only Polish companies but also foreign ones. This will allow to compare the results and highlight potential differences in the supplier evaluation criteria. In addition, it will increase the representativeness of the research and hence the results will be more universal.

9 Acknowledgments

The paper was written with the financial support from the National Center of Science the grant of the no. DEC-2011/03/D/HS4/03367.

References

1. Ageron, B., Gunasekaran, A., and Spalanzani, A. (2013) "IS/IT as supplier selection criterion for upstream value chain." *Industrial Management and Data Systems* 113 (3):443-460.
2. Al-Tabtabai H. and Diekmann, J. (1992) "Judgemental forecasting in construction projects." *Construction Management and Economics* 10 (1):19-30.
3. Anholcer, M. and Kawa, A. (2012) "Optimization of supply chain via reduction of complaints ratio." In *Agent and Multi-Agent Systems. Technologies and Applications*, edited by Jezic, G. et al., pp. 622-628. Springer, Berlin Heidelberg.
4. Bosstorff, P. and Rosenbaum, R. G. (2007) "Supply chain excellence: a handbook for dramatic improvement using the SCOR model (Vol. 1)." Amacom.
5. Burton, T. (1988) "JIT repetitive sourcing strategies: tying the knot with your suppliers." *Production and Inventory Management Journal* 29 (4):38-41.
6. Chen, H., Daugherty, P. J. and Roath, A. S. (2009). "Defining and operationalizing supply chain process integration." *Journal of Business Logistics* 30 (1):63-84.
7. Ching-Lai, H. and Kwangsun, Y. (1981) "Multiply Attribute Decision Making. Methods and Applications." Springer, Berlin Heidelberg.
8. Choi, T. Y. and Krause, D. R. (2006) "The supply base and its complexity: Implications for transaction costs, risks, responsiveness, and innovation." *Journal of Operations Management* 24 (5):637-652.
9. Cousins, P. D. (1999) "Supply base rationalisation: myth or reality?" *European Journal of Purchasing and Supply Management* 5 (3):143-155.
10. Crowley, L.G., et al. (1995) "Evaluation of Competitive Bids." *Journal of Construction Engineering and Management* 121 (2):238-243.
11. Duszak, Z. and Koczkodaj, W.W. (1994) "The Generalization of a New Definition of Consistency for Pairwise Comparisons." *Information Processing Letters* 52:273-276.

12. Fulop, J. Koczkodaj, W.W. and Szarek, S. (2010) "A different perspective on a scale for pairwise comparisons." In *Transactions on computational collective intelligence I*, edited by Nguyen N.T and Kowalczyk R., 71-84. Springer, Berlin Heidelberg.
13. Herbon, A., Moalem, S., Shnaiderman, H. and Templeman, J. (2012) "Dynamic weights approach for off-line sequencing of supplier selection over a finite planning horizon." *International Journal of Physical Distribution and Logistics Management*, 42 (5):434-463.
14. Herman, M. and Koczkodaj, W.W. (1996) "Monte Carlo Study of Pairwise Comparisons." *Information Processing Letters* 57 (1):25-29.
15. Hofmann, E. and Kotzab, H. (2010) "A supply chain-oriented approach of working capital management." *Journal of Business Logistics* 31 (2):305-330.
16. Igoulalene, I., Benyoucef, L. and Tiwari, M. K. (2015). Novel fuzzy hybrid multi-criteria group decision making approaches for the strategic supplier selection problem. *Expert Systems with Applications*, 42(7), 3342-3356.
17. Ishizaka, A. (2014). Comparison of fuzzy logic, AHP, FAHP and hybrid fuzzy AHP for new supplier selection and its performance analysis. *International Journal of Integrated Supply Management*, 9(1-2), 1-22.
18. Jensen, R.E. (1984) "An Alternative Scaling Method for Priorities in Hierarchical Structures", *Journal of Mathematical Psychology* 28:317-332.
19. Kawa, A. (2012). "SMART logistics chain." In *Intelligent Information and Database Systems*, pp. 432-438, Springer Berlin Heidelberg.
20. Koczkodaj, W.W., Kulakowski, K., Ligeza, A. (2014) "On the quality evaluation of scientific entities in Poland supported by consistency-driven pairwise comparisons method." *Scientometrics* 99(3):911-926.
21. Koczkodaj, W.W. Szwarc, R. (2014) "On Axiomatization of Inconsistency Indicators for Pairwise Comparisons", *Fundamenta Informatica* <http://arxiv.org/abs/1307.6272>
22. Koczkodaj, W.W. (1993) "A New Definition of Consistency of Pairwise Comparisons." *Mathematical and Computer Modelling* 18 (7):79-84.
23. Koczkodaj, W.W. (1996) "Statistically Accurate Evidence of Improved Error Rate by Pairwise Comparisons." *Perceptual and Motor Skills* 82:43-48.
24. Leenders, M., Johnson, P.F., Flynn, A. and Fearon, H.E. (2006) "Purchasing Supply Management", 13th ed., McGraw-Hill, New York.
25. Lin, C., Kou, G., and Ergu, D. (2013) "A statistical approach to measure the consistency level of the pairwise comparison matrix." *Journal of the Operational Research Society* 65 (9):1380-1386.
26. Memon, M. S., Lee, Y. H. and Mari, S. I. (2015). Group multi-criteria supplier selection using combined grey systems theory and uncertainty theory. *Expert Systems with Applications*, 42(21), 79517959. 1993.
27. Monczka, R.M., Trent, R. and Handfield, R. (2002) "Purchasing and Supply Chain Management." 2nd ed., South-Western/Thomson Learning, Cincinnati, OH.
28. Nijkamp, P., Rietveld, P. and Voogd, H. (1990) "Multicriteria Evaluation in Physical Planning." North-Holland, Amsterdam.
29. Ogden, J. A. (2006) "Supply base reduction: an empirical study of critical success factors." *Journal of Supply Chain Management* 42 (4):29-39.
30. Sarkar, A. and Mohapatra, P. K. (2006) "Evaluation of supplier capability and performance: A method for supply base reduction". *Journal of Purchasing and Supply Management* 12 (3):148-163.
31. Sim, H. A. (1956) "Rational Choice and the Structure of the Environment". *Psychological Review* 63 (2):129-138.
32. Supply Chain Council (2010) "SCOR, Supply Chain Operations Reference Model, ver. 10.0." USA.
33. Tully, S. (1995) "Purchasing's new muscle." *Fortune* 131 (3):75-84.
34. Thurstone, L.L. (1927) "A Law of Comparative Judgments." *Psychological Reviews* 34:273-286.
35. Vokurka, R. J., Choobineh and J., Vadi, L. (1996) "A prototype expert system for the evaluation and selection of potential suppliers." *International Journal of Operations and Production Management* 16 (12):106-127.

36. Voogd, H. (1983) "Multicriteria Evaluation for Urban and regional Planning." Pion Ltd., London.
37. Whitten, J. and Bentley, L. (2005) "Systems analysis and design methods" McGraw-Hill Professional.
38. Zacharia, Z. G., Nix, N. W. and Lusch, R. F. (2009) "An analysis of supply chain collaborations and their effect on performance outcomes." *Journal of Business Logistics* 30 (2):101-123.
39. Zhang, X., Deng, Y., Chan, F. T. and Mahadevan, S. (2015). A fuzzy extended analytic network process-based approach for global supplier selection. *Applied Intelligence*, 1-13.
40. Zouggar, A. and Benyoucef, L. (2012) "Simulation based fuzzy TOPSIS approach for group multi-criteria supplier selection problem." *Engineering Applications of Artificial Intelligence* 25(3):507-519.