

Sergio Zani, Maria Adele Milioli, and Isabella Morlini

Abstract

Composite indicators should ideally measure multidimensional concepts which cannot be captured by a single variable. In this chapter, we suggest a method based on fuzzy set theory for the construction of a fuzzy synthetic index of a latent phenomenon (e.g., well-being, quality of life, etc.), using a set of manifest variables measured on different scales (quantitative, ordinal and binary). A few criteria for assigning values to the membership function are discussed, as well as criteria for defining the weights of the variables. For ordinal variables, we propose a fuzzy quantification method based on the sampling cumulative function and a weighting system which takes into account the relative frequency of each category. An application regarding the results of a survey on the users of a contact center is presented.

Keywords

Imprecise data and fuzzy methods • Membership function • Overall satisfaction • Quantification of ordinal variables • Weighting criteria

23.1 Introduction

The purpose of composite indicators is to measure multidimensional concepts which cannot be captured by a single variable. They are usually formed on the basis of a set of quantitative variables [11, 16]. In order to construct a composite indicator, in this chapter we suggest a method based on the fuzzy set theory [9, 17, 18, 21]. As shown in the literature, fuzzy sets permit the representation of vague concepts, e.g., poverty [5, 15], well-being [2, 7], quality of life [14], business scenarios [4],

S. Zani (✉) · M.A. Milioli · I. Morlini
Department of Economics, University of Parma, Italy
e-mail: sergio.zani@unipr.it; milioli@unipr.it; isabella.morlini@unimore.it

satisfaction indices [20], etc. In Sect. 23.2, we deal with the general problem of obtaining a synthetic fuzzy measure of a latent phenomenon from a set of manifest variables measured on different scales (quantitative, ordinal, and binary). We present a few criteria to transform the values of a quantitative variable into fuzzy numbers. For ordinal variables we propose a fuzzy quantification method based on the cumulative function, and for a set of binary variables we consider the relative frequency of the symptoms of the underlying concept. In Sect. 23.3, we discuss the problem of weighting the variables and aggregating them in a composite indicator. The weights should reflect the contribution of each variable to the latent phenomenon. For ordinal variables, we suggest a new criterion taking into account the relative frequency of each category. In Sect. 23.4, we focus on the specific application of measuring customer satisfaction using ordinal (and binary) variables. The gradual transition from very dissatisfied to really satisfied customers can be captured by a fuzzy index. We apply the suggested methods to a sample of 704 respondents of a survey on the users of a contact center, in order to evaluate their satisfaction with the service. The fuzzy indicator of customer satisfaction allows us to obtain a ranking of respondents that can be compared with the traditional ones. Furthermore, the comparison between the overall satisfaction scores directly stated by the respondents and the values of the fuzzy composite indicator based on several items shows noncoherent answers for a few customers (high stated satisfaction but low value of the synthetic index, or viceversa). These units may be considered as atypical observations.

23.2 Fuzzy Transformations of the Variables

Let X be a set of elements $x \in X$. A fuzzy subset A of X is a set of ordered pairs:

$$[x, \mu_A(x)] \quad \forall x \in X \quad (23.1)$$

where $\mu_A(x)$ is the membership function (m.f.) of x to A in the closed interval $[0,1]$. If $\mu_A(x) = 0$, then x does not belong to A , while if $\mu_A(x) = 1$ then x completely belongs to A . If $0 < \mu_A(x) < 1$, then x partially belongs to A and its membership to A increases according to the values of $\mu_A(x)$. Let us assume that the subset A defines the position of each element with reference to achievement of the latent concept, e.g., the well-being of a set of countries or the satisfaction of a sample of customers. In this case, $\mu_A(x) = 1$ identifies a situation of full achievement of the target (a country enjoying global well-being or a completely satisfied customer), whereas $\mu_A(x) = 0$ denotes a total failure (a country with no well-being or a very dissatisfied customer).

Consider a set of n units or elements e_i ($i = 1, 2, \dots, n$) and p manifest variables X_s ($s = 1, 2, \dots, p$) reflecting the latent phenomenon. Without loss of generality, let us assume that each variable is positively related to that phenomenon, i.e., it satisfies the property “the larger the better.” If a quantitative variable X_s shows negative correlation, we substitute it with a simple decreasing function transformation, e.g.,

$f(x_{si}) = \max(x_{si}) - x_{si}$. In case of an ordinal variable, we consider it in reverse order. In order to define the m.f. for each variable it is necessary:

1. To identify the extreme situation such that $\mu_A(x) = 0$ (non-membership) and $\mu_A(x) = 1$ (full membership).
2. To define a criterion for assigning m.f. values to the intermediate categories of the variable.

Let us assume that X_s is a quantitative variable; for simplicity of notation we omit index s . For that variable X , we choose an inferior (lower) threshold l and a superior (upper) threshold u , with l and u finite, and we define the m.f. as follows:

$$\begin{cases} \mu_A(x_i) = 0, & x_i \leq l \\ \mu_A(x_i) = \frac{x_i - l}{u - l}, & l < x_i < u \\ \mu_A(x_i) = 1, & x_i \geq u \end{cases} \quad (23.2)$$

In (23.2) m.f. is a linear function between the values of the two thresholds. Alternatively, we can arrange the values x_i in nondecreasing order according to i and define the following m.f.:

$$\begin{cases} \mu_A(x_i) = 0, & x_i \leq l \\ \mu_A(x_i) = \mu_A(x_{i-1}) + \frac{F(x_i) - F(x_{i-1})}{1 - F[x_{i(l)}]}, & l < x_i < u \\ \mu_A(x_i) = 1, & x_i \geq u \end{cases} \quad (23.3)$$

where $F(x_{si})$ is the sampling cumulative function of the variable X and $x_{i(l)}$ is the highest value $x_i \leq l$. If $l = x_1 = \min(x_i)$ and $u = x_n = \max(x_i)$, formula (23.3) corresponds to the “totally fuzzy and relative approach” suggested by Cheli and Lemmi [6]. In the literature, other specifications have been considered. If we are trying to measure the degree of achievement of a certain target, the distance $d(x)$ between the value x and the goal is an indicator of the success in achieving the target. If $d(x) = 0$, there is full membership to A , then $\mu_A(x) = 1$. If $d(x) > 0$ then $\mu_A(x) < 1$. Hence, we can write:

$$\mu_A(x) = \frac{1}{1 + d(x)} \quad (23.4)$$

In general, as highlighted by Zimmermann [21], the relationship between physical measures and perception takes an exponential form. The distance $d(x)$ can be expressed as:

$$d(x) = e^{-a(x-b)} \quad (23.5)$$

so that m.f. is defined as follows:

$$\mu_A(x) = \frac{1}{1 + e^{-a(x-b)}} \quad (23.6)$$

Table 23.1 Membership function to the subset of satisfied customers with reference to variable X_8

Scores (x_i)	n_i	$F(x_i)$	m.f.(x_i)
≤ 3	7	0.010	0
4	5	0.017	0
5	12	0.034	0
6	40	0.091	0.059
7	161	0.320	0.296
8	240	0.732	0.723
9	93	0.864	1
10	96	1	1
	704		

It is worth noting that the parameter a represents the extent of vagueness and the parameter b may be viewed as the point in which the tendency of the subject’s attitude changes from rather positive to rather negative. Balamoune-Lutz [1] uses m.f. (23.6) to measure human well-being with a fuzzy approach. If variable X_s is ordinal with k categories, a suitable choice is the following: the m.f. values of the categories up to the threshold l are equal to 0 (absence of the phenomenon) and those of the categories $\geq u$ are equal to 1 (complete presence). The intermediate values of m.f. are defined according to the formula (23.3). For example, consider the scores (on a scale 1–10) of the variable “Overall satisfaction,” described in Sect. 23.4. Choosing $l = 5$ and $u = 9$, we obtain the m.f. values indicated in the last column of Table 23.1. A customer with a score equal to 5 does not belong to the set of satisfied respondents, while a customer with a score equal to 8 has a value 0.723 of the m.f. to the set of satisfied users. If variable X_s is binary, one of the categories can be considered as a symptom of the latent concept. Therefore the m.f. is a crisp function assuming only values equal to 0 (absence) and 1 (presence). Usually, we consider a set of q ($q \leq p$) binary variables reflecting several aspects of the phenomenon. In this case the m.f. can be defined as follows:

$$\mu_A(x_i) = \frac{1}{q} \sum_{s=1}^q z_{si} \tag{23.7}$$

where $z_{si} = 1$ if the corresponding x_{si} denotes presence of the symptom and $z_{si} = 0$ if the corresponding x_{si} denotes absence of the symptom. Definition (23.7) is consistent with interpreting membership values as the proportion of “subjects” who rate the i -th unit as an actual member of the fuzzy set A [5].

23.3 Weighting and Aggregation of the Variables

Among the steps of the construction of a crisp composite indicator, weighting and aggregation are the major ones which directly affect the quality and reliability of the results [16, 22]. For the aggregation functions in fuzzy set theory we refer, among others, to Calvo and Beliakov [3].

Let us consider the criteria for aggregating the p fuzzy variables, described in Sect. 23.2, into a fuzzy composite indicator. The simplest operations for the i -th unit are:

$$\text{intersection : } \quad I \mu_A(i) = \min[\mu_A(x_{1i}), \mu_A(x_{2i}), \dots, \mu_A(x_{pi})] \quad (23.8)$$

$$\text{union : } \quad U \mu_A(i) = \max[\mu_A(x_{1i}), \mu_A(x_{2i}), \dots, \mu_A(x_{pi})] \quad (23.9)$$

A more general aggregation function is the weighted generalized means [12]:

$$\mu_A(i) = \sum_{s=1}^p [\mu_A(x_{si})]^\alpha \cdot w_s^{1/\alpha} \quad (23.10)$$

where $w_s > 0$ is the normalized weight that expresses the relative importance of the variables X_s ; ($\sum_{s=1}^p w_s = 1$). For fixed arguments and weights, function (23.10) is monotonic increasing with α ; if $\alpha \rightarrow -\infty$, then formula (23.10) becomes the intersection defined in (23.8); if $\alpha \rightarrow +\infty$, then (23.10) is equal to the union (23.9). For $\alpha \rightarrow 0$ formula (23.10) becomes the weighted geometric mean. In the following, for the sake of simplicity, we will consider $\alpha = 1$, that is the weighted arithmetic mean. The weighting criteria in (23.10) may be:

- Equal weights that imply a careful selection of the variables in order to assure a balance of the different aspects of the latent phenomenon.
- Factor loadings, obtained by principal components analysis (PCA) for quantitative variables or by nonlinear PCA for ordinal variables; this method of weighting is valid if the first component accounts for a high percentage of the total variance and the weights (loadings) of the variables are proportional to their correlation with the first component (factor) reflecting the underlying concept.
- Obtained from expert judgements, e.g., using focus groups.
- Determined by an Analytic Hierarchy Process [13].

We suggest a criterion for the determination of the weights, considering for each variable X_s the fuzzy proportion $g(X_s)$ of the achievement of the target:

$$g(X_s) = \frac{1}{n} \sum_{i=1}^n \mu(x_{si}) \quad (23.11)$$

If X_s is binary, formula (23.11) coincides with the crisp proportion and in general it may be seen as an index of the proportion of the units having (totally or partially) the latent phenomenon [6]. The normalized weights may be determined as an inverse function of $g(X_s)$, in order to give higher importance to the rare features in the n units. To avoid excessive weights to the variables with low value of $g(X_s)$ we choose [5]:

$$w_s = \ln \left[\frac{1}{g(X_s)} \right] / \sum_{s=1}^p \ln \left[\frac{1}{g(X_s)} \right] \quad (23.12)$$

Using (23.12), it is possible to attach to each variable a weight sensitive to the fuzzy membership.

23.4 A Fuzzy Indicator of Customer Satisfaction

Customer satisfaction may be defined as the degree of happiness that a customer experiences with a product or a service and is a personal function of the gap between expected and perceived quality. The expected degree of happiness is a random vague concept. This latent phenomenon is here considered for all individuals and analyzed with fuzzy method [8, 10]. We apply the methods of the previous sections to the results of a survey on customer satisfaction of the users of the Contact Center of the Municipality of Parma, considering a sample of 704 respondents calling for information (see [19] for a complete description). The questions on the degree of satisfaction of the users are:

- X_1 = Contact at the first call (no = 0, yes = 1).
- X_2 = Waiting time (too long = 1, normal = 2, fairly short = 3).
- X_3 = Courtesy of the operator.
- X_4 = Skill of the operator.
- X_5 = Quality of the provided information.
- X_6 = Speed of the information.
- X_7 = Complete answer (no = 1, partially = 2, yes = 3).
- X_8 = Overall satisfaction with the service, with scores from 1 to 10.

All the variables whose categories are not specified are measured on Likert scale (very dissatisfied = 1, dissatisfied = 2, neither satisfied nor dissatisfied = 3, satisfied = 4, very satisfied = 5). The cumulative function of the variables $X_1 - X_7$ and the corresponding m.f. according to formula (23.3) are presented in Table 23.2 and for variable X_8 in the previous Table 23.1. For the variables on Likert scale, the inferior threshold is “dissatisfied” and the superior threshold is “very satisfied.” We consider the following weighting systems, with and without the variable X_8 , which indicates overall satisfaction with the service (see Table 23.3):

1. Equal weights of the variables (W_1).
2. Normalized factor loadings obtained by standard (linear) PCA. The first PC accounts for 49.7 % of the total variance with X_8 and for 50.4 % without X_8 (W_2).
3. Normalized factor loadings applying PCA on τ rank correlation matrix. The first PC explains 46.67 % of the total variance with X_8 and 46.44 % without X_8 (W_3).
4. Normalized weights as inverse function of the fuzzy proportion of each variable, according to formula (23.12) (W_4).

The least important variable is always X_1 and could be deleted. The correlation coefficient between W_2 and W_3 is 0.964 considering X_8 and 0.970 without X_8 , but the correlation coefficients of W_4 with the previous weights are in the interval [0.791, 0.909]. Therefore, the last weighting criterion is slightly different from the others.

Table 23.2 Cumulative function F of the categories of each item in the sample and corresponding membership function to the subset A of satisfied customers

X_1	X_2		X_3		X_4		V_5		X_6		X_7									
	F	$m.f.$	F	$m.f.$	F	$m.f.$	F	$m.f.$	F	$m.f.$	F	$m.f.$								
0	0.08	0	1	0.03	0	1	0	0	1	0.01	0	1	0.01	0	1	0.05	0			
1	1	1	2	0.29	0.27	2	0.01	0	2	0.03	0	2	0.04	0	2	0.04	0			
			3	1	1	3	0.05	0.05	3	0.14	0.12	3	0.13	0.09	3	0.12	0.08	3	1	1
						4	0.46	0.45	4	0.60	0.59	4	0.57	0.55	4	0.58	0.56			
						5	1	1	5	1	1	5	1	1	5	1	1			

Table 23.3 Values of W_2 , W_3 and W_4

	With X_8			Without X_8		
	W_2	W_3	W_4	W_2	W_3	W_4
X_1	6.97	6.93	3.59	7.99	7.83	4.38
X_2	9.21	8.84	10.57	10.62	10.05	12.88
X_3	12.29	13.67	13.65	14.17	15.64	16.63
X_4	14.76	15.50	16.18	16.95	17.71	19.71
X_5	15.61	16.08	16.43	17.97	18.25	20.02
X_6	17.73	16.02	16.22	18.13	18.32	19.76
X_7	12.27	10.67	5.43	14.17	12.20	6.62
X_8	13.16	12.29	17.93	–	–	–
Tot	100	100	100	100	100	100

Table 23.4 Frequency distribution of the fuzzy composite indicators with different weights

Classes	With X_8				Without X_8			
	W_1	W_2	W_3	W_4	W_1	W_2	W_3	W_4
0.0┆0.1	4	7	7	8	5	7	7	7
0.1┆0.2	8	12	12	12	7	12	12	14
0.2┆0.3	14	13	16	20	15	15	15	18
0.3┆0.4	36	36	37	39	30	34	33	35
0.4┆0.5	30	31	33	26	35	28	29	30
0.5┆0.6	49	49	46	120	26	27	37	81
0.6┆0.7	109	158	151	103	91	205	195	142
0.7┆0.8	134	96	106	82	161	68	68	69
0.8┆0.9	105	77	66	99	122	69	69	66
0.9┆1.0	129	136	144	109	42	69	69	72
1.0	86	86	86	86	170	170	170	170
Total	704	704	704	704	704	704	704	704

Table 23.4 shows the frequency distribution of the values of the fuzzy composite indicators with the mentioned weighting criteria. None of the respondents can be regarded as completely dissatisfied, since the values of the composite indicators are at least 0.02. Even clients experiencing dissatisfaction with most indicators

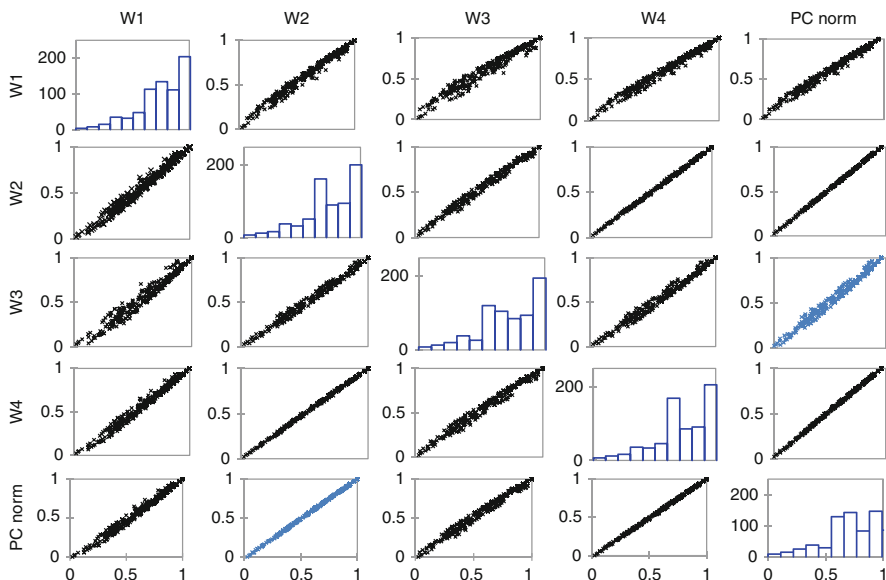


Fig. 23.1 Scatter plot of the fuzzy indicators (obtained with inclusion of X_8) and of the crisp indicator (obtained normalizing the scores of the first principal component)

are found to be not completely dissatisfied with other indicators. On the other hand, 86 respondents (with X_8) and 170 (without X_8) belong to the subset of completely satisfied customers. In both analyses, the use of the weights increases the frequencies of small values (less than 0.4) and decreases the frequencies in the classes $[0.7 - 0.8)$ and $[0.8 - 0.9)$.

Figure 23.1 displays the scatter plot matrix of the fuzzy indicators obtained with inclusion of X_8 and of the crisp indicator obtained normalizing the scores of the first principal component. As outlined previously, fuzzy and crisp indices show high pairwise correlations (the correlation coefficient of each off-diagonal panel of the scatter plot matrix is nearly 0.99). With $l = 1$ the distribution of the values is more symmetric. In order to compare the fuzzy indicators computed without X_8 with the mostly used crisp indicators of customer satisfaction, we normalize both the values of X_8 (overall satisfaction) and the scores of the first principal component computed on the τ rank correlation matrix, to lie in the interval $[0, 1]$. Figure 23.2 presents boxplots of the fuzzy and crisp indices. While all composite indicators show the presence of outlying values (due to small size), the single variable X_8 does not reveal this presence. Moreover, all fuzzy indicators show that 75% of the respondents belong to the set of satisfied customers, with m.f. values higher than 0.6. Figure 23.3 shows the distribution of the fuzzy index (computed without X_8 and with weight W_4) for each category of the variable X_8 . Note that none of the respondents has a category smaller than 3 (i.e., no one is completely dissatisfied with the service). Boxplots show that the distribution of the values of the m.f. to the subset

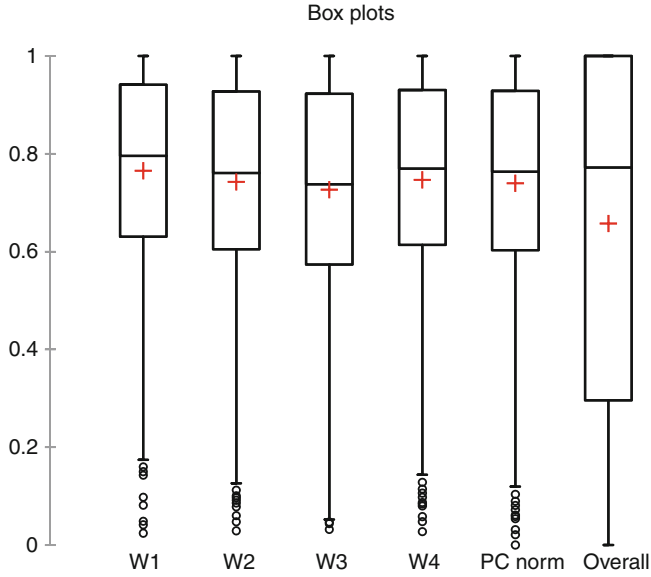


Fig. 23.2 Boxplots of the fuzzy indicators computed without X_8 , of the normalized values of variable X_8 (Overall) and of the scores of the first principal component (Pcnorm)

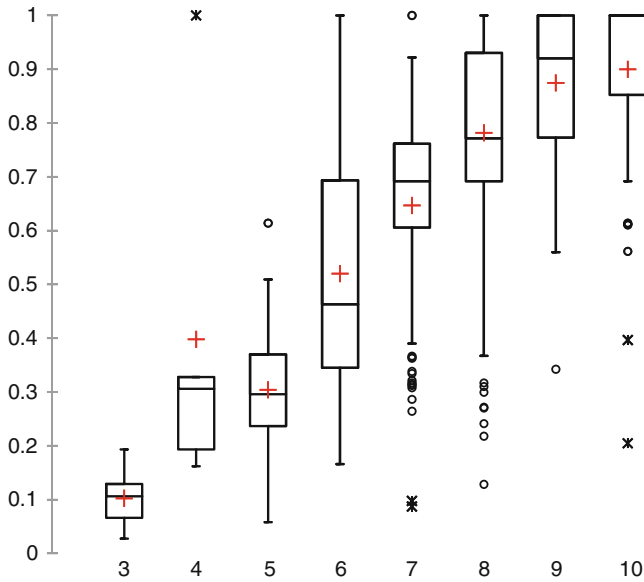


Fig. 23.3 Boxplots of the fuzzy indicators with weights W_4 for each value of the variable X_8 (overall satisfaction for the service)

of satisfied customers is increasing in the median, with respect to the score given to the overall satisfaction (X_8). It also reveals the presence of few incoherent responses. For example, two customers that have a score equal to 4 in the overall satisfaction (they are neither satisfied nor very dissatisfied) are however quite satisfied with all the specific items that have been considered in the questionnaire. On the other hand, some clients that have a scores equal to 8, 9, or 10 in the overall satisfaction indicate a low degree of satisfaction in nearly all the items. We point out that the comparison between the direct question on overall satisfaction and the values of a fuzzy index based on the items reveals the customers with inconsistent answers who may be considered as atypical observations.

References

1. Balamoune-Lutz, M.: On the measurement of human well-being: Fuzzy set theory and Sen's capability approach. In: UNU-WIDER, Helsinki (2004)
2. Balamoune-Lutz, M., McGillivray, M.: Fuzzy well-being achievement in Pacific Asia. *J. Asia Pacific Econ.* **11**, 168–177 (2006)
3. Calvo, T., Beliakov, G.: Aggregation functions based on penalties. *Fuzzy Sets Syst.* **161**, 1420–1436 (2010)
4. Castillo, C., Lorenzana, T.: Evaluation of business scenarios by means of composite indicators. *Fuzzy Econ. Rev.* **15**, 3–20 (2010)
5. Cerioli, A., Zani, S.: A fuzzy approach to the measurement of poverty. In: Dagum, C., Zenga, M. (eds.) *Income and Wealth Distribution, Inequality and Poverty*, pp. 272–284. Springer, Berlin (1990)
6. Cheli, B., Lemmi, A.: A totally fuzzy and relative approach to the multidimensional analysis of poverty. *Econ. Notes* **24**(1), 115–134 (1995)
7. Chiappero Martinetti, E.: A multidimensional assessment of well-being based on Sen's functioning approach. *Rivista Internazionale di Scienze Sociali* **108**(2), 207–239 (2000)
8. Chien, C.-J., Tsai, H.-H.: Using fuzzy numbers to evaluate perceived service quality. *Fuzzy Sets Syst.* **116**, 289–300 (2000)
9. Coppi, R., Gil, M.A., Kiers, H.A.L.: A fuzzy approach to statistical analysis. *Comput. Stat. Data Anal.* **51**, 1–14 (2006)
10. Darestani, A.Y., Jahromi, A.E.: Measuring customer satisfaction using a fuzzy inference system. *J. Appl. Sci.* **9**(3), 469–478 (2009)
11. JRC: An information server on composite indicators and ranking systems. <http://composite-indicators.jrc.ec.europa.eu>. (2013)
12. Klir, G.J., Folger, T.A.: *Fuzzy Sets, Uncertainty and Information*, p. 61. Prentice-Hall Int., London (1988)
13. Kwong, C.K., Bai, H.: A fuzzy AHP approach to the determination of importance weights of customer requirements in quality function deployment. *J. Intell. Manufact.* **13**, 367–377 (2002)
14. Lazim, M.A., Osman, M.T.A.: A new Malaysian quality of life index based on Fuzzy sets and hierarchical needs. *Soc. Indicat. Res.* **94**(3), 499–508 (2009)
15. Lemmi, A., Betti, G. (eds.): *Fuzzy Set Approach to Multidimensional Poverty Measurement*. Springer, New York (2006)
16. OCDE: *Handbook on Constructing Composite Indicators*. OCDE Publications, Paris (2008)
17. Smithson, M., Verkuilen, J.: *Fuzzy Sets Theory: Applications in the Social Sciences*. Sage, London (2006)
18. Zadeh, L.A.: Fuzzy sets. *Inform. Contr.* **8**, 338–353 (1965)
19. Zani, S., Berziera, L.: Measuring customer satisfaction using ordinal variables: an application in a survey on a contact center. *Stat. Applicata-Italian J. Appl. Stat.* **20**(3–4), 331–351 (2008)

20. Zani, S., Milioli, M.A., Morlini, I.: Fuzzy methods and satisfaction indices. In: Kenett, R.S., Salini, S. (eds.) *Modern Analysis of Customer Surveys*, pp. 439–455. Wiley, Chichester (2012)
21. Zimmermann, H.J.: *Fuzzy Sets Theory and Its Applications*, 4th edn. Kluwer, Boston (2001)
22. Zhou, P., Ang, B.W., Zhou, D.Q.: Weigthing and aggregation in composite indicator construction: a multiplicative optimization approach. *Soc. Indicat. Res.* **96**, 169–181 (2010)