

Sustainability performance evaluation in industry by composite sustainability index

Li Zhou · Hella Tokos · Damjan Krajnc ·
Yongrong Yang

Received: 22 October 2011 / Accepted: 20 January 2012 / Published online: 9 February 2012
© Springer-Verlag 2012

Abstract The growing importance of sustainable development as a policy objective has initiated a debate about those suitable frameworks and tools useful for policy makers when making a sustainable decision. Composite indicators (CIs) aggregate multidimensional issues into one index, thus providing comprehensive information. However, it is frequently argued that CIs are too subjective, as their results undesirably depend on the normalization method, a specific weighting scheme, and the aggregation method of sub-indicators. This article applies different combinations of normalization, weighting, and aggregation methods for the assessment of an industrial case study, with the aim of determining the best scheme for constructing CIs. The applied methodology gradually aggregates sustainable development indicators into sustainability sub-indices and, finally, to a composite sustainability index. The normalization methods included in this analysis are: minimum–maximum, distance to a reference, and the percentages of annual differences over consecutive years. Equal weightings, the ‘benefit of the doubt’ approach, and budget allocation process were used for determining the weights of individual indicators and sustainability sub-indices. The linear, geometric, and non-compensatory multi-criteria approaches (NCMCs) were used as aggregation methods. The NCMC is

modified to fit the two-level aggregation, then to sub-indices, and finally to a composite sustainable index. Also, a penalty criterion is introduced into the evaluation process with the aim of motivating the company to move towards sustainable development. The results are analyzed by variance-based sensitivity analysis. According to the results the recommended scheme for CIs’ construction is: distance to a reference–benefit of the doubt–linear aggregation.

Keywords Sustainability assessment · Composite sustainability index · Sensitivity analysis · Breweries

Abbreviations

AHP	Analytical hierarchy processes
BAP	Budget allocation process
BAT	Best available techniques
BOD	Benefit of the doubt
CA	Conjoint analysis
CI	Composite indicator
DEA	Data envelopment analysis
EW	Equal weighting
FAST	Fourier amplitude sensitivity test
GME	Geometric aggregation
GRI	Global reporting initiative
LIN	Linear aggregation
MCDA	Multiple criteria decision analysis
NCMC	Non-compensatory multi-criteria approach
UCM	Unobserved component models
WP	Weighted product

List of symbols

Sets

I Set of indicator or input quantity of the evaluated model

L. Zhou · H. Tokos · Y. Yang (✉)
State Key Laboratory of Chemical Engineering, Department
of Chemical and Biochemical Engineering, Zhejiang University,
Hangzhou, Zhejiang 310027, People’s Republic of China
e-mail: yangyr@zju.edu.cn

D. Krajnc (✉)
Laboratory for Process Systems Engineering and Sustainable
Development, Faculty of Chemistry and Chemical Engineering,
University of Maribor, Smetanova 17, 2000 Maribor, Slovenia
e-mail: damjan.krajnc@uni-mb.si

- J Group of indicators (group of environmental indicators: $j = 1$; group of societal indicators: $j = 2$; group of economic indicators: $j = 3$) or input quantity of the evaluated model
- T The analysed time period (2003–2007)

Parameters

- $I_{i,j,t}^+$ Indicator i from group of indicator j in year t with positive impact on sustainable development
- $I_{i,j,t}^-$ Indicator i from group of indicator j in year t with negative impact on sustainable development
- $I_{i,j}^{\text{Benchmark}}$ Benchmark for indicator i from group of indicators j
- $I_{i,j}^{+, \text{MAX}}$ The highest value of indicator i with positive impact on sustainable development from group of indicator j for the analyzed time period
- $I_{i,j}^{-, \text{MAX}}$ The highest value of indicator i with negative impact on sustainable development from group of indicator j for the analyzed time period
- $I_{i,j}^{+, \text{MIN}}$ The lowest value of indicator i with positive impact on sustainable development from group of indicator j for the analyzed time period
- $I_{i,j}^{-, \text{MIN}}$ The lowest value of indicator i with negative impact on sustainable development from group of indicator j for the analyzed time period
- X_i Input quantity i of the evaluated model
- \tilde{X}_i The true value of input quantity i

Variables

- $E(YX_i)$ Expected value for the output quantity Y for the whole variation interval of the input quantity X_i
- $E[V(Y|X_{-i})]$ Expected amount of residual variance when X_i , and only X_i were left free to vary over its uncertainty range, all the other variables are fixed
- $I_{N_{i,j,t}}^+$ Normalized indicator i from group of indicator j in year t with positive impact on sustainable development
- $I_{N_{i,j,t}}^-$ Normalized indicator i from group of indicator j in year t with negative impact on sustainable development
- $I_{S_{j,t}}$ Sustainable sub-indices for group of indicator j in year t
- $I_{S_{j,t}}^{\text{Benchmark}}$ Sustainable sub-indices for group of indicator j determined for the benchmarks in year t
- $I_{S_{j,t}}^*$ The highest value for the sustainable sub-indices for group of indicator j in year t
- $I_{\text{SUST}, t}$ Composite sustainability index in year t

- $I_{\text{SUST}, t}^*$ The highest value for the composite sustainability index in year t
- S_i First-order sensitivity index for input quantity i
- $S_{i,j}$ Second-order sensitivity index or two-way interaction for input quantities X_i and X_j
- S_{Ti} Total sensitivity index for input quantity i
- $V(Y)$ Unconditional variance for the output variable Y
- $V[E(Y|\tilde{X}_i)]$ Conditional variance of the expected value for the output quantity Y when the input quantity is fixed on its true value \tilde{X}_i
- $V[E(Y|X_i, X_j)]$ Conditional variance of the expected value for the output quantity Y when input quantities X_i and X_j are fixed
- w_j Weight of the group of sustainability indicator (sub-indices) j
- $w_{i,j}$ Weight of indicator i from group of indicator j
- Y Output quantity of the evaluated model

Introduction

Policy makers within industry are paying increased attention to implementing the sustainable development concept into business activities due to fierce competition in the global market, and strict environmental regulations. Sustainable development indicators are recognized as useful tools for the assessment and anticipation of production performance and trends, being able to provide early-warning information to prevent economic, societal, and environmental damage, and support decision-making (Singh et al. 2009). The indicators can be divided in two groups: content indicators, which are describing the state of the system and performance indicators, which are measuring the behavior of the system (Sikdar 2003). The composite indicator (CI) can be defined as an aggregation of different indicators according to a well-developed and pre-determined methodology (Gasparatos et al. 2008). CIs can be divided into several different categories depending on the various methods selected during their formulation (Niemeijer 2002): (1) data-driven, when data availability is the central issue concerning the development of the CIs and high-quality data must be provided, (2) theory-driven, when selecting the best possible indicators for CIs construction is done from a theoretical point of view, and data availability is only one of the many aspects considered, (3) policy-driven, when the indicators are selected, especially for the monitoring of a certain policy. Various

methodologies exist for the construction of CIs. Nardo et al. (2008) described a framework for the construction of a CI, which includes the selection of relevant indicators and data, imputation of missing data, normalization of the selected indicators, weighting, and aggregation. Even though all steps are important for the quality of the final CI, the weighting and aggregation steps seem to have the greatest impact. Zhou et al. (2007) proposed a mathematical-programming approach for the construction of CIs using multiple criteria decision analysis (MCDA) for aggregation of the selected indicators and two sets of weights (the most and least favorable) that are generated from the data themselves. This proposed approach does not require a prior knowledge in relation to the weights of the selected indicators. Cherchye (2007) used data envelopment analysis (DEA) in their construction with the aim of neutralizing the recurring sources of criticism about CIs. The application of this method makes it possible to skip the normalization stage as it is invariant to measurement units and the weights are generated by a 'flexible benefit of the doubt' method. Sikdar (2009) simplified the comparison process of aggregated metrics by expressed them as the geometric mean of the ratios of the individual metrics for pairwise comparison. The generated aggregate metric by this method is sensitive to the individual metrics and weighting factors, also. Hatefi and Torabi (2010) proposed a common weight MCDA–DEA approach for constructing CIs. Zhou et al. (2010) proposed a multiplicative optimization approach for constructing CIs, using the weighted product (WP) method. In their approach the weights are generated by solving a series of multiplicative DEA-type models that can be transformed into equivalent linear programs. The developed model enables the incorporation of additional relevant information about the weights.

The sustainability assessment is performed in several manufacturing sectors, such as the steel industry (Singh et al. 2007), chemical industry (Beloff and Tanzil 2006), traditional beet-sugar plants (Krajnc et al. 2007), dry-cleaning industry (von Bahr et al. 2003), and breweries (Tokos et al. 2011). There have also been studies that focused on benchmarking the energy efficiency of an industry (Phylipsen et al. 2002; Mateos-Espejel et al. 2011; Frangopoulos and Keramioti 2010), transport sector (Henning et al. 2011), sludge-plants (Abusam et al. 2004), water management (De Carvalho et al. 2009), and ski areas (Geneletti 2008).

CIs remain controversial despite their increasing usage. It is frequently argued that CIs are too subjective, due to the assumptions when estimating any measurement error in data, the selection of the relevant indicators, choice of normalization scheme, weights and aggregation systems (Singh et al. 2009). It is vital to find the best combination for a normalization–weighting–aggregation scheme

regarding the construction of a CI, which will effectively measure any changes in a company's performance, and assist it toward a more sustainable development.

By using sensitivity analysis, it can be determined how the composite sustainability index depends on the information fed into it. Chan et al. (1997) applied and compared several variance-based methods (correlation ratios or importance measures, Sobol' indices, and Fourier amplitude sensitivity test (FAST) indices), in order to measure how much a model depends on its input. They concluded that the correlation-ratio measurement is model-independent, and can only evaluate the first-order sensitivity indices of the input parameters. 'FAST and Sobol' are completely automated and are able to compute those total-effect indices that allow for a qualitative ranking of the input parameters in regard to their influence on the output. Saltelli (2002) introduced a new strategy for the computation of full sets of first and total-effect sensitivity indices for a model's output. The computation of sensitivity indices are based on decomposing the variance of the target-function in a quantitative manner. Cherchye et al. (2006) applied uncertainty and sensitivity analyses to assess the robustness of the final results and to evaluate the influences of each individual uncertainty source on the output variance, if DEA was used for the construction of CI. Saisana and Saltelli (2008), using information collected from budget allocation (BAP) and analytical hierarchy processes (AHPs), analyzed the relative importance of expert opinion on the selected indicators included in a CI.

This article, aims to identify the best integration scheme for the construction of a composite sustainability index, via a sensitivity analysis of its application during an industrial case study. The remainder of this article is organized as follows: steps for the construction of a composite suitability index and the considered methodologies, together with the used sensitivity analysis, are given in second section. The industrial case study is presented in third section, followed by the results and conclusions in fourth section.

Construction of a composite sustainability index, and sensitivity analysis

By applying sensitivity analysis to the obtained results, the objective is to identify the more suitable methods for constructing a composite sustainability index. The evaluated methodology gradually aggregates the selected environmental, societal, and economic indicators into a composite sustainability index, Fig. 1. The relevant indicators covering different aspects of sustainability are selected during the pre-modeling stage, and their positive or negative impacts on the sustainability development are

judged. These steps are followed by data collection, including indicators, benchmarks, and weights. The modeling phase involves normalization that transforms the indicators into dimensionless forms. Weights are assigned to the individual indicators and sustainability sub-indices before the aggregation process, thus measuring their importance for sustainable development within the company. By applying a step-by-step procedure, the indicators are first grouped into sustainability sub-indices, which are then combined into a composite sustainability index. At the end of the methodological procedure, the final result provides a set of sub-indices for different sustainability dimensions in regard to the company's performance, thus providing a degree of sustainability implementation.

Indicator selection, judgment on the indicator's impact, and data collection

According to the described methodology, the first step during the integrated sustainability performance assessment is the selection of a suitable set of indicators. In our approach, the indicators were selected based on global reporting initiative (GRI) guidelines. These guidelines provide a set of core and additional indicators. The core indicators are those indicators, which are identified as interesting to most stakeholders and assumed to be material, unless deemed otherwise, on the basis of the GRI reporting principles. However, the final selection should always be performed in close cooperation with the industry. Any additional indicators are optional, for example, some additional indicators can be included if the company is interested in evaluating its influence on sustainable development. Likewise, some core indicators can also be excluded, as they are not being measured or are seen as unimportant. The environmental indicators are selected

from three aspects: (1) Material, (2) Energy, and (3) Emissions, Effluents, Waste. The societal indicators are chosen from the next three groups: (1) Employment, (2) Occupational Health and Safety, and (3) Diversity and Equal Opportunities. Eight economic indicators are presented from the aspect of economic performance.

In the next step, the indicators are divided into groups in regard to their influences on sustainable development. These groups consist of those whose increasing value have a positive impact on sustainable development (I^+), e.g., paper recycling, and alternatively those whose increasing values have a negative impact (I^-), e.g., freshwater consumption. Depending on which group the indicator is assigned to, the normalization equation will be different for every normalization method applied.

Data collection is the most time-consuming process, as a large number of values need to be collected from the company regarding their environmental, societal, and economic performances. In addition, the benchmarks need to be defined, and are later used during the normalization method (distance to reference). The 'benefit of the doubt' approach is used for determining the weights. The benchmark values were determined for each of the selected indicators, based on the values from the best available techniques (BAT), the measurements and standards within the company, the local legal regulations, GRI reports for specific production sectors, and other relevant documents. The strictest limit within the range is selected for those cases where the benchmark values are given within a range. Typical types of technology, the raw materials, the production process, and the types of energy used during production, are taken into account when determining the benchmarks. Expert opinion regarding the importance of individual indicators and sustainability sub-indices also needs to be collected for the calculation of weights within BAP.

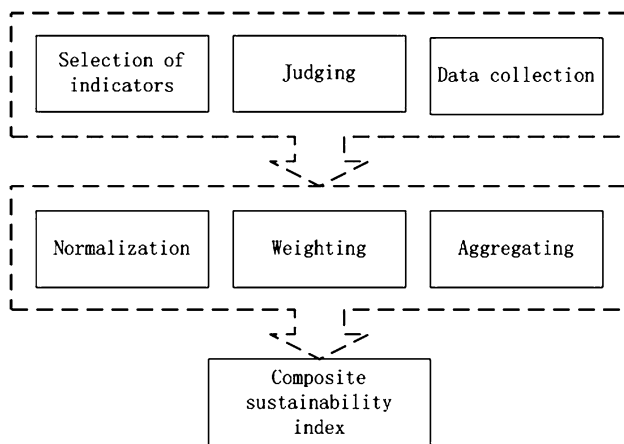


Fig. 1 Scheme for calculation of composite sustainability index

Normalization

Normalization is necessary for integrating the selected indicators into a composite sustainability index, as they are usually expressed in different units. The normalization methods included are: minimum–maximum, distance to a reference, and the percentage of annual differences over consecutive years.

Minimum–Maximum

According to this method, each indicator with a positive impact on sustainable development, $I_{i,j,t}^+$, is transformed into a normalized form by the equation:

$$I_{N_{i,j,t}}^+ = \frac{I_{i,j,t}^+ - I_{i,j}^{+,MIN}}{I_{i,j}^{+,MAX} - I_{i,j}^{+,MIN}} \quad \forall i \in I \quad I_{i,j}^{+,MAX} = \max_{t \in T} I_{i,j,t}^+ \wedge$$

$$I_{i,j}^{+,MIN} = \min_{t \in T} I_{i,j,t}^+ \quad (1)$$

While the indicator with a negative impact on sustainable development, $I_{i,j,t}^-$, is normalized by the equation:

$$I_{N_{i,j,t}}^- = 1 - \frac{I_{i,j,t}^- - I_{i,j}^{-,MIN}}{I_{i,j}^{-,MAX} - I_{i,j}^{-,MIN}} \quad \forall i \in I \quad I_{i,j}^{-,MAX} = \max_{t \in T} I_{i,j,t}^- \wedge$$

$$I_{i,j}^{-,MIN} = \min_{t \in T} I_{i,j,t}^- \quad (2)$$

where $I_{i,j,t}^+$ and $I_{i,j,t}^-$ are the values for indicator i from the group of indicator j in year t with positive and negative impacts on sustainable development, respectively, while $I_{N_{i,j,t}}^+$ and $I_{N_{i,j,t}}^-$ are their normalized indicators, respectively. The highest value for indicator i with positive impact on sustainable development from the group of indicator j for the analyzed time period is denoted as $I_{i,j}^{+,MAX}$, while for indicator i with negative impact on sustainable development, as $I_{i,j}^{-,MAX}$. Otherwise, the lowest value for indicator i with positive and negative impact are denoted as $I_{i,j}^{+,MIN}$, and $I_{i,j}^{-,MIN}$, respectively. In this way the normalized indicator with positive impact on sustainable development will have a value between 0, for $I_{i,j,t}^+ = I_{i,j}^{+,MIN}$, and 1, for $I_{i,j,t}^+ = I_{i,j}^{+,MAX}$. In the case of an indicator with negative impact on sustainable development, the highest value will be achieved when $I_{i,j,t}^- = I_{i,j}^{-,MIN}$, according to the indicator's nature. It should be noted, that this transformation is time-dependent, which implies an adjustment of the analyzed time period, if new data is available. This adjustment may change the minimum and maximum values of some indicators, and then affect the normalized values. The composite sustainability index for the existing data must be re-calculated to maintain comparability between the existing and the new data.

Distance to a reference

When applying this method, the normalized value is calculated as the ratio between the indicator and an external benchmark. The external benchmark can be defined by the BAT values, measurements and standards for a specific production sector, local legal regulations, GRI reports, and any other relevant documents. The normalized indicators are described by Eqs. 3 and 4:

$$I_{N_{i,j,t}}^+ = \frac{I_{i,j,t}^+}{I_{i,j}^{Benchmark}} \quad (3)$$

$$I_{N_{i,j,t}}^- = \frac{I_{i,j}^{Benchmark}}{I_{i,j,t}^-} \quad (4)$$

where $I_{i,j}^{Benchmark}$ is the benchmark for indicator i from the group of indicators j . Using the denominator $I_{i,j}^{Benchmark}$, the equation takes into account the evolution of indices over time, according to possible future benchmark updates. In this case, the normalized value can have a value higher than 1, indicating that the company has above-average performance.

Percentage of annual differences over consecutive years

The indicators are transformed by Eqs. 5 and 6:

$$I_{N_{i,j,t}}^+ = \frac{I_{i,j,t}^+ - I_{i,j,t-1}^+}{I_{i,j,t-1}^+} \quad (5)$$

$$I_{N_{i,j,t}}^- = \frac{I_{i,j,t-1}^- - I_{i,j,t}^-}{I_{i,j,t-1}^-} \quad (6)$$

The disadvantage of this method is that the indicators for $t = t^0$ cannot be normalized with the given equations, and would be lost during the analysis.

Weighting methods

The relative importance of indicators is a source of disagreement, as the decision makers of companies have different views, and are interested in different indicators. The weights of indicators can be obtained by statistical models, such as factor analysis, DEA, and unobserved component models (UCM), or from participatory methods such as, BAPs, AHPs, and conjoint analysis (CA). The applied weighting methods are presented in the continuation.

Equal weightings (EWs)

Most of the CIs are constructed by EW, which means that all indicators are assigned the same weight. This essentially implies that all indicators have the same importance, but it could also disguise the absence of statistical or empirical bases for determining the weights. In any case, EW does not mean no weights, but implicitly implies that the weights are equal. Moreover, if variables are grouped into dimensions and these are further aggregated into the composite, then applying EW to the variables may imply an unequal weighting of the dimensions (the dimensions grouping the larger number of variables will have higher weight). This could result in an imbalanced structure within the composite index.

'Benefit of the doubt' approach (BOD)

Using this method, the composite sustainability index in year t , $I_{SUST,t}$, is defined as the ratio between the actual performance of the company and the external benchmark:

$$I_{\text{SUST}_t} = \frac{\sum_j I_{S_{j,t}} \cdot w_j}{\sum_j I_{S_j}^{\text{Benchmark}} \cdot w_j} \quad (7)$$

where $I_{S_{j,t}}$ is the sustainable sub-indices for the group of indicator j in year t , while $I_{S_{j,t}}^{\text{Benchmark}}$ is their benchmarks, and w_j is the weight of the group of the sustainability indicator (sub-indices) j , reflecting the importance given to the environmental, societal, and economic performance of the company (group of environmental indicators: $j = 1$; group of societal indicators: $j = 2$; group of economic indicators: $j = 3$). The sustainable sub-indices, $I_{S_{j,t}}$, are calculated by the equation:

$$I_{S_{j,t}} = \sum_i I_{N_{i,j,t}}^+ \cdot w_{i,j} + \sum_i I_{N_{i,j,t}}^- \cdot w_{i,j}$$

$$\sum_i w_{i,j} = 1, \quad (8)$$

$$w_{i,j} \geq 0$$

While $I_{S_{j,t}}^{\text{Benchmark}}$ is given by:

$$I_{S_{j,t}}^{\text{Benchmark}} = \sum_i I_{i,j}^{\text{Benchmark}} \cdot w_{i,j}$$

$$\sum_i w_{i,j} = 1, \quad (9)$$

$$w_{i,j} \geq 0$$

where $w_{i,j}$ is the weight of indicator i from the group of indicator j , and reflects the importance of this indicator during the sustainability assessment of the company. With the aim of determining the weights for individual indicators, a linear programming problem can be defined, which ensures the highest value for the CI:

$$I_{S_{j,t}}^* = \arg \max_{w_{i,j}} \left(\sum_i I_{N_{i,j,t}}^+ \cdot w_{i,j} + \sum_i I_{N_{i,j,t}}^- \cdot w_{i,j} \right)$$

$$\sum_i I_{i,j}^{\text{Benchmark}} \cdot w_{i,j} \leq 1, \quad (10)$$

$$w_{i,j} \geq 0$$

where $I_{S_{j,t}}^*$ is the highest value for the sustainable sub-indices for the group of indicator j in year t . The weight of the group of sustainability indicator, w_j , can be determined in a similar way by:

$$I_{\text{SUST}_t}^* = \arg \max_{w_j} \left(\sum_j I_{S_{j,t}} \cdot w_j \right)$$

$$\sum_j I_{S_j}^{\text{Benchmark}} \cdot w_j \leq 1, \quad (11)$$

$$w_j \geq 0$$

where $I_{\text{SUST}_t}^*$ is the highest value for the composite sustainability index in year t .

Budget allocation process (BAP)

The BAP determines the indicator weights based on expert opinion. In order to establish a proper weighting system, it is essential to bring together experts representing a wide-spectrum of knowledge and experience, e.g., experts from different production sectors and management within the company. They estimate a preference factor for each indicator from the aspect of sustainability by following a scale from 1 (unimportant) to 5 (highly important), based on their experience and subjective judgment of the selected indicators' relative importance. The weights are then calculated as average preference factors. The main advantages of BAP are its transparent and relatively straightforward nature, and its short duration. The disadvantage of this method is that the weights could reflect specific local conditions, e.g., environmental problems, or the need for political intervention regarding some indicators, e.g., local legal regulations regarding CO₂ emission, so the weighting may not be transferable from one region to another.

Aggregation methods

When applying a step-by-step procedure the selected indicators are grouped into sustainability sub-indices. The $I_{S_{j,t}}$, for each group of sustainability indicators j , are then combined into a composite sustainability index, I_{SUST_t} . The aggregation methods used during this analysis were: linear, geometric, and non-compensatory multi-criteria approaches (NCMCs).

Linear aggregation (LIN)

As Eq. 8 illustrated, the sustainability sub-indices for the group of indicator j , $I_{S_{j,t}}$ were calculated as a summation of weighted and normalized individual indicators. Based on this, when applying LIN, the composite sustainability index is calculated as:

$$I_{\text{SUST}_t} = \sum_j I_{S_{j,t}} \cdot w_j$$

$$\sum_j w_j = 1, \quad (12)$$

$$w_j \geq 0$$

LIN is widely used, because of its simplicity, transparency, and easy understanding. An undesirable feature of this method is its compensability, namely the poor performances of some indicators can be compensated by the sufficiently high values of other indicators thus, in this way, the composite sustainability index will not entirely reflect any information about its individual indicators.

Geometric aggregation (GME)

The shortcomings of the LIN method can be partially overcome by using GME, where the sustainability sub-indices for the group of indicator j , $I_{S_{j,t}}$, is calculated as a product of the normalized individual indicators as a power of their weights:

$$I_{S_{j,t}} = \prod_i \left(I_{N_{ij,t}}^+ \right)^{w_{ij}} \cdot \prod_i \left(I_{N_{ij,t}}^- \right)^{w_{ij}} \quad (13)$$

and the composite sustainability index is calculated by the equation:

$$I_{\text{SUST}_t} = \prod_j \left(I_{S_{j,t}} \right)^{w_j} \quad (14)$$

Use of GMEs is recommended when non-comparable and strictly positive sub-indicators are expressed in different ratio scales. It should be noticed that compensability exists in case of GMEs, also. In case of LINs the compensability is constant, while for GME is partial, i.e., compensability is lower when the CI contains indicators with low values. In other words, GME is able to alleviate but not to eliminate the compensability. Moreover, as a method of multiplication, problems may occur when zero is present in the input data.

Non-compensatory multi-criteria approach (NCMC)

In the cases of linear and GMEs, the substitution rates among the indicators are equal to the weights of the indicators. As a consequence, the weights in these aggregation schemes necessarily have the meaning of substitution rates and do not indicate the importance of the associated indicator. When using the NCMC during the construction of a composite sustainability index, the weights would be interpreted as important coefficients. The NCMC approach can be divided into two steps: (1) pairwise comparison of indicators with the whole set of selected indicators for different years, (2) ranking of years based on an achieved sustainability performance, which would give information about the company's progress towards sustainable

production. These steps are first applied for the ranking of sustainability sub-indices (environmental, societal, and economic), and then repeated during the second step using the results from the previous ranking, in order to grade the composite sustainability index for all $t \in T$. As an illustrative example, three environmental indicators (freshwater, heat, and electricity consumption), two societal (total number of employees and the number of injuries), and two economic indicators (profit and operating costs) were selected. In addition, a set of weights, considered as important coefficients and which satisfied the condition $\sum_i w_{i,j} = 1$ for every sustainability sub-indices group j , were assumed for an individual indicator. The data necessary for constructing an outranking impact matrix for sustainability sub-indices, is given in the impact matrix, Table 1. The score for each year t within the environmental sub-indices is the sum of those weights of individual indicators that showed better performance in year t . Table 2 gives the outranking impact matrix for environmental sub-indices. For the year 2003, in comparison with 2004, the score is equal to 0, because during the year 2003 the company had higher water, heat, and electricity consumption than in 2004. Accordingly, the score for 2004 in comparison to 2003 is equal to 1. It should be noted that environmental indicators have a negative impact on sustainable development, therefore, a lower value is preferable. The final ranking would be the permutation with the highest score. If two or more permutations have the same score, the permutation from the earliest year would be ranked higher as a penalty for the following years, when no improvement was achieved in the field of sustainable development. The outranking impact matrixes for societal and economic sub-indices are shown in Tables 3 and 4, respectively. It should be stressed that, in the case of societal indicators, the total number of employees was positive, while the number of injuries had a negative impact on sustainable development. In the same manner, profit was positive while the operating cost had a negative impact on sustainable development among the selected economic indicators. The ranks determined for the

Table 1 Impact matrix for the illustrative example—first level of aggregation

	Environmental indicator			Societal indicator		Economical indicator	
	Freshwater (m ³ /a)	Heat (MJ/a)	Electricity (GW h/a)	No. of employees	No. of injuries	Profit (M€/a)	Operating cost (M€/a)
2003	750,319	151,469	15,995	423	20	99,347	80,968
2004	636,572	133,663	13,831	381	19	89,245	72,193
2005	653,305	124,669	13,055	367	15	80,885	68,212
2006	715,512	127,434	13,586	339	24	87,047	69,186
2007	636,250	127,253	13,928	331	14	124,320	94,356
Weight	2/4	1/4	1/4	1/2	1/2	1/3	2/3

Table 2 Outranking impact matrix for environmental sub-indices

	Environmental sub-indices					Rank
	2003	2004	2005	2006	2007	
2003	0	0	0	0	0	5
2004	1	0	2/4	2/4	1/4	3
2005	1	2/4	0	1	2/4	1
2006	1	2/4	0	0	1/4	4
2007	1	3/4	2/4	3/4	0	2

Table 3 Outranking impact matrix for societal sub-indices

	Societal sub-indices					Rank
	2003	2004	2005	2006	2007	
2003	0	0.5	0.5	1	0.5	1
2004	0.5	0	0.5	1	0.5	2
2005	0.5	0.5	0	1	0.5	3
2006	0	0	0	0	0.5	5
2007	0.5	0.5	0.5	0.5	0	4

sustainability sub-indices are joined into a new impact matrix on the second level of aggregation, Table 5, and equal weights were assigned to the sub-indices. According to the outranking impact matrix for the composite sustainability index (Table 6), the company had the best sustainability performance in the year 2005. With the aim of being able to compare the results of NCMC with the results of LIN and GME, the rankings of the composite sustainability index were divided by the highest possible permutation score, thus placing the results within a range of between 0 and 1. The highest possible permutation score for the illustrative example is 4.

Sensitivity analysis

The development of a CI involves stages where subjective judgments have to be made: when selecting a suitable set of indicators, the choice of the normalization method, the choice of the aggregation method, the weights of the indicators, etc. By using sensitivity analysis, it can be determined as to how the variation in the output is connected qualitatively or quantitatively, to the different sources of variation within the assumptions, and how the composite sustainability index depends on the information fed into it. It was noted earlier that CIs may be considered as models. When several layers of uncertainty are present simultaneously, a CI could become a non-linear, possibly non-additive model. In the case of sensitive analysis regarding non-linear models, variance-based techniques should be used, which are model-free and robust techniques. Our

Table 4 Outranking impact matrix for economic sub-indices

	Economic sub-indices					Rank
	2003	2004	2005	2006	2007	
2003	0	1/3	1/3	1/3	2/3	4
2004	2/3	0	1/3	1/3	2/3	3
2005	2/3	2/3	0	2/3	2/3	1
2006	2/3	2/3	1/3	0	2/3	2
2007	1/3	1/3	1/3	2/3	0	5

Table 5 Impact matrix—second level of aggregation

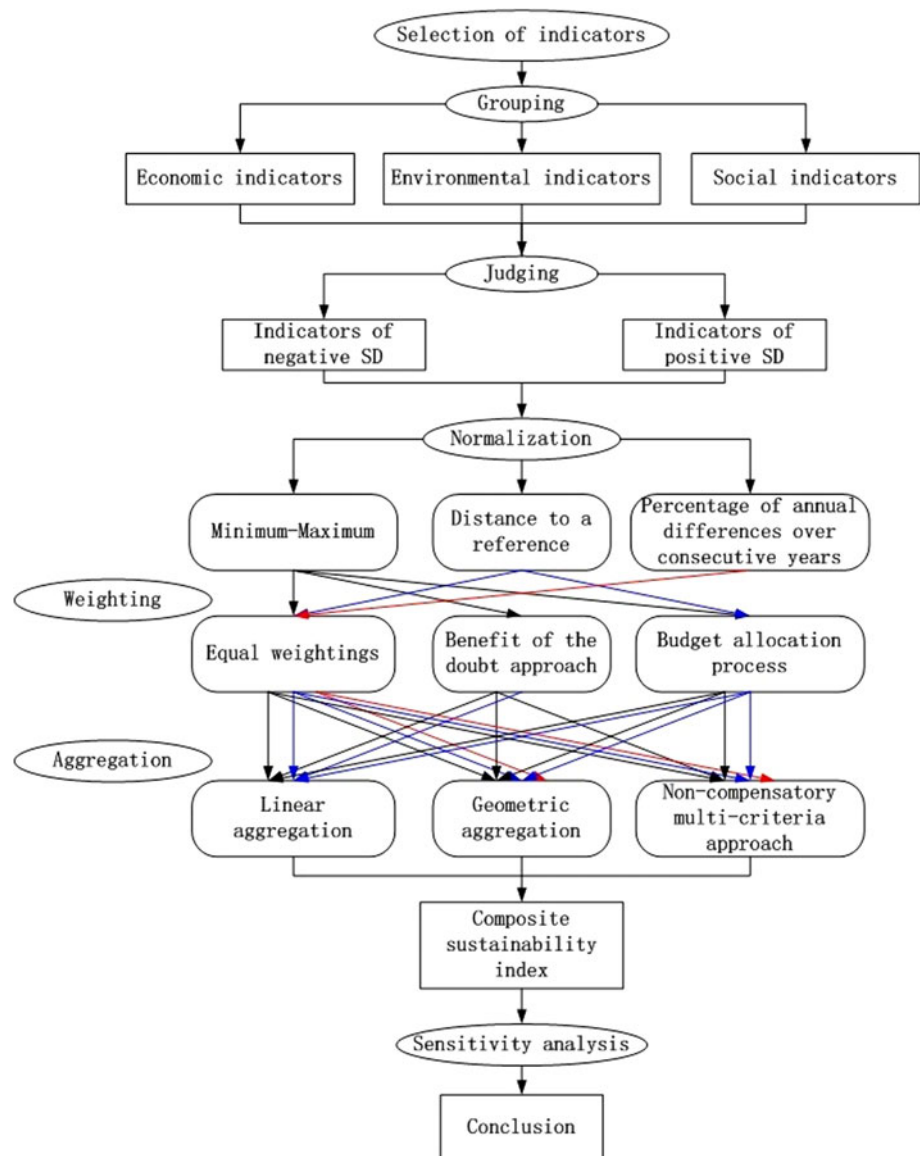
	Sustainability sub-indices		
	Environmental	Societal	Economic
2003	5	1	4
2004	3	2	3
2005	1	3	1
2006	4	5	2
2007	2	4	5
Weight	1/3	1/3	1/3

Table 6 Outranking impact matrix for composite sustainability index

	Composite sustainability index					Rank	I_{SUST}
	2003	2004	2005	2006	2007		
2003	0	1/3	1/3	1/3	2/3	3	0.415
2004	2/3	0	1/3	2/3	2/3	2	0.582
2005	2/3	2/3	0	1	1	1	0.832
2006	2/3	1/3	0	0	1/3	4	0.332
2007	1/3	1/3	0	1/3	0	5	0.250

analysis mainly focused on three uncertainties: different normalization methods, different weighting schemes, and different aggregation methods. The evaluation procedure for the composite sustainability index model is given in Fig. 2. This procedure involves three normalization methods, three weighting schemes, and three aggregation methods, thus creating a total of 19 combinations for the calculation of a composite sustainability index. In variance-based sensitivity analysis, it is assumed that the true value, \tilde{X}_i , of an input quantity X_i is known. A conditional variance $V[E(Y|\tilde{X}_i)]$ is estimated holding the true value of the specific fixed-input quantity. Unfortunately, in general, the true values of the input quantities are unknown. Therefore, in order to obtain global sensitivity measurements, the expected value $E(Y|X_i)$ above the whole variation interval of the input quantity X_i has to be evaluated. Variance-based sensitivity indices are estimated as ratios between the

Fig. 2 Evaluation procedure for the composite sustainability index model



conditional variance and the unconditional variance, $V(Y)$, for the output quantity Y :

$$S_i = \frac{V[E(Y|X_i)]}{V(Y)} \quad (15)$$

where S_i is the first-order sensitivity index. This sensitivity index indicates the relative importance of an individual input quantity X_i when driving the uncertainty. For additive models, it holds true that $\sum_i S_i = 1$, which means that the input is non-correlated. This leads to an easy quantitative interpretation of the sensitivity index, because each S_i delivers a direct measurement for that portion of X_i on the output variance $V(Y)$. However, in the cases of non-additive models, it is necessary to take into consideration the interactions among the input quantities within the models or the effects of the higher order. The terms of the

higher order are estimated by holding more than one fixed-input quantity, e.g., the input quantities X_i and X_j :

$$S_{i,j} = \frac{V[E(Y|X_i, X_j)]}{V(Y)} - S_i - S_j \quad (16)$$

where $S_{i,j}$ is the second-order sensitivity index or a two-way interaction for input quantities X_i and X_j and $V[E(Y|X_i, X_j)]$ is the conditional variance of the expected value for the output quantity Y when the input quantities X_i and X_j are fixed. The computations of all higher-order terms have high computational costs. However, estimation of the total effects, S_{Ti} , includes all higher-order terms with respect to the input quantity X_i , Eq. 17, makes it possible to calculate within one computational step, like the first-order sensitivity index:

$$S_{Ti} = \frac{E[V(Y|X_{-i})]}{V(Y)} \tag{17}$$

where $E[V(Y|X_{-i})]$ is the expected amount of residual variance if X_i , and only X_i were left free to vary over its uncertainty range, all the other variables having their true values (they are fixed). While the first-order sensitivity indices are quantitative sensitivity measurements for additive models, the total sensitivity indices are quantitative measurements for all kinds of models, independent of their model characteristics. Comparison between S_i and S_{Ti} may lead to a conclusion regarding the additivity of a model with non-correlated input, as for additive models $S_{Ti} = S_i$ and for non-additive models $S_{Ti} > S_i$. Additionally, the difference $S_{Ti} - S_i$ is a measure of how much X_i is involved in any interaction with other input variables.

Case study

Selecting the best combination from the analyzed methodologies described in “Construction of a composite sustainability index, and sensitivity analysis” section was carried out as a case study of a brewery. The sustainable indicators were chosen according to the GRI guidelines in cooperation with the company, and grouped under three sections covering the economic, environmental, and social dimensions of sustainability. In total, 69 environmental, 22 social, and 8 economic performance indicators were selected, which were delivered for the time period 2003–2007. The selected environmental, societal and economic indicators are the same as in Tokos et al. (2011). The weights of individual indicators and groups of indicators determined by EW, BAP, and BOD, are given in Table 7 for the environmental group of indicators, in Table 8 for the societal, and in Table 9 for the economic group of indicators. The weights for the BAP method were determined based on the opinion of 11 experts from different production sectors within the case-studied brewery. The defined benchmarks for the selected indicators and more details regarding the analyzed case study can be found in Tokos et al. (2011). The sensitivity analysis was performed over the results of 19 combinations of methods for constructing a composite sustainability index.

Results and discussion

The results from uncertainty analysis of the composite sustainability index for different construction combinations are given in Fig. 3. It can be concluded that:

Table 7 Weights of individual environmental indicators and sub-index

	Indicators			Sub-index		
	EW	BOD	BAP	EW	BOD	BAP
Aspect: Materials				0.333	0.8	0.333
$\psi_{\text{water/beer}}$	0.0145	0.0001	0.132			
$m_{\text{malt}}^{\text{Brewhouse1}}/V_{\text{beer}}$	0.0145	0.0001	0.124			
$m_{\text{maize}}^{\text{Brewhouse1}}/V_{\text{beer}}$	0.0145	0.0001	0.124			
$m_{\text{hops}}^{\text{Brewhouse1}}/V_{\text{beer}}$	0.0145	0.0001	0.124			
$m_{\text{malt}}^{\text{Brewhouse2}}/V_{\text{beer}}$	0.0145	0.0001	0.124			
$m_{\text{maize}}^{\text{Brewhouse2}}/V_{\text{beer}}$	0.0145	0.0001	0.124			
$m_{\text{hops}}^{\text{Brewhouse2}}/V_{\text{beer}}$	0.0145	0.0001	0.124			
$m_{\text{dia.earth}}/V_{\text{beer}}$	0.0145	0.0001	0.124			
Aspect: Energy						
Q/V_{beer}	0.0145	0.00068	0.482			
E/V_{beer}	0.0145	0.0001	0.518			
Aspect: Emissions, effluents, and waste						
$m_{\text{CO}_2}^{\text{Boiler}}/V_{\text{beer}}$	0.0145	0.0001	0.019			
$m_{\text{CO}_2}^{\text{Electricity}}/V_{\text{beer}}$	0.0145	0.0001	0.019			
$m_{\text{CO}}^{\text{Boiler}}/V_{\text{beer}}$	0.0145	0.0001	0.019			
$m_{\text{NO}_x}^{\text{Boiler}}/V_{\text{beer}}$	0.0145	0.993	0.019			
$m_{\text{CO}}^{\text{Electricity}}/V_{\text{beer}}$	0.0145	0.0001	0.019			
$m_{\text{NO}_x}^{\text{Electricity}}/V_{\text{beer}}$	0.0145	0.0001	0.019			
$m_{\text{dust}}/V_{\text{beer}}$	0.0145	0.0001	0.019			
$V_{\text{ww}}^{\text{Brewhouse}}/V_{\text{beer}}$	0.0145	0.0001	0.021			
$c_{\text{pH}}^{\text{Brewhouse}}/V_{\text{beer}}$	0.0145	0.0001	0.021			
$c_{\text{i.m.}}^{\text{Brewhouse}}/V_{\text{beer}}$	0.0145	0.0001	0.021			
$c_{\text{s.m.}}^{\text{Brewhouse}}/V_{\text{beer}}$	0.0145	0.0001	0.021			
$c_{\text{COD}}^{\text{Brewhouse}}/V_{\text{beer}}$	0.0145	0.0001	0.021			
$c_{\text{BOD}}^{\text{Brewhouse}}/V_{\text{beer}}$	0.0145	0.0001	0.021			
$c_{\text{Cu}}^{\text{Brewhouse}}/V_{\text{beer}}$	0.0145	0.0001	0.021			
$c_{\text{AOX}}^{\text{Brewhouse}}/V_{\text{beer}}$	0.0145	0.0001	0.021			
$c_{\text{P}}^{\text{Brewhouse}}/V_{\text{beer}}$	0.0145	0.0001	0.021			
$c_{\text{N}}^{\text{Brewhouse}}/V_{\text{beer}}$	0.0145	0.0001	0.021			
$c_{\text{NH}_4\text{-N}}^{\text{Brewhouse}}/V_{\text{beer}}$	0.0145	0.0001	0.021			
$c_{\text{Zn}}^{\text{Brewhouse}}/V_{\text{beer}}$	0.0145	0.0001	0.021			
$c_{\text{TCl}}^{\text{Brewhouse}}/V_{\text{beer}}$	0.0145	0.0001	0.021			
$c_{\text{Cl}}^{\text{Brewhouse}}/V_{\text{beer}}$	0.0145	0.0001	0.021			
$c_{\text{Cl}^-}^{\text{Brewhouse}}/V_{\text{beer}}$	0.0145	0.0001	0.021			
$c_{\text{TOC}}^{\text{Brewhouse}}/V_{\text{beer}}$	0.0145	0.0001	0.021			
$V_{\text{ww}}^{\text{Cooling}}/V_{\text{beer}}$	0.0145	0.0001	0.017			
$c_{\text{pH}}^{\text{Cooling}}/V_{\text{beer}}$	0.0145	0.0001	0.017			
$c_{\text{i.m.}}^{\text{Cooling}}/V_{\text{beer}}$	0.0145	0.0001	0.017			
$c_{\text{s.m.}}^{\text{Cooling}}/V_{\text{beer}}$	0.0145	0.0001	0.017			
$c_{\text{COD}}^{\text{Cooling}}/V_{\text{beer}}$	0.0145	0.0001	0.017			
$c_{\text{BOD}}^{\text{Cooling}}/V_{\text{beer}}$	0.0145	0.0001	0.017			

Table 7 continued

	Indicators			Sub-index		
	EW	BOD	BAP	EW	BOD	BAP
$C_{w.flea}^{Cooling}/V_{beer}$	0.0145	0.0001	0.017			
$C_{Cu}^{Cooling}/V_{beer}$	0.0145	0.0001	0.017			
$C_{AOX}^{Cooling}/V_{beer}$	0.0145	0.0001	0.017			
$C_P^{Cooling}/V_{beer}$	0.0145	0.0001	0.017			
$C_{N-NO_2}^{Cooling}/V_{beer}$	0.0145	0.0001	0.017			
$C_{Zn}^{Cooling}/V_{beer}$	0.0145	0.0001	0.017			
$C_{TCl}^{Cooling}/V_{beer}$	0.0145	0.0001	0.017			
$C_{Cr}^{Cooling}/V_{beer}$	0.0145	0.0001	0.017			
$C_{HC}^{Cooling}/V_{beer}$	0.0145	0.0001	0.017			
V_{ww}^{Boiler}/V_{beer}	0.0145	0.0001	0.021			
C_{pH}^{Boiler}/V_{beer}	0.0145	0.0001	0.021			
$C_{i.m.}^{Boiler}/V_{beer}$	0.0145	0.0001	0.021			
$C_{s.m.}^{Boiler}/V_{beer}$	0.0145	0.0001	0.021			
$C_{COD}^{Boiler}/V_{beer}$	0.0145	0.0001	0.021			
$C_{BOD}^{Boiler}/V_{beer}$	0.0145	0.0001	0.021			
C_{Pb}^{Boiler}/V_{beer}	0.0145	0.0001	0.021			
$C_{AOX}^{Boiler}/V_{beer}$	0.0145	0.0001	0.021			
C_P^{Boiler}/V_{beer}	0.0145	0.0001	0.021			
$C_{N-NO_2}^{Boiler}/V_{beer}$	0.0145	0.0001	0.021			
C_{Fe}^{Boiler}/V_{beer}	0.0145	0.0001	0.021			
$C_{NH_4-N_2}^{Boiler}/V_{beer}$	0.0145	0.0001	0.021			
$C_{TOC}^{Boiler}/V_{beer}$	0.0145	0.0001	0.021			
C_{HC}^{Boiler}/V_{beer}	0.0145	0.0001	0.021			
Waste by type and disposal method						
m_{glass}/V_{beer}	0.0145	0.0001	0.143			
m_{paper}/V_{beer}	0.0145	0.0001	0.143			
$m_{plastic}/V_{beer}$	0.0145	0.0001	0.143			
m_{metal}/V_{beer}	0.0145	0.0001	0.143			
m_{grain}/V_{beer}	0.0145	0.0001	0.143			
m_{yeast}/V_{beer}	0.0145	0.0001	0.143			
$m_{dia.earth}/V_{beer}$	0.0145	0.0001	0.143			

- the composite sustainability index value varied from 0.0 to 0.9, as well as the ranking order of each year indicating that different combinations of the selected methodologies can lead to very different results, and a comparison of different methods is a necessity,
- as long as the GME method was involved within the construction frame, the composite sustainability index had zero value when there was zero value in the former data layer.

By detailed analysis of the data layer, it was found that the zero value had two sources: (1) origin data collected

Table 8 Weights of individual societal indicators and sub-index

	Indicators			Sub-index		
	EW	BOD	BAP	EW	BOD	BAP
Aspect: Employment				0.333	0.1	0.319
$N_{manag.}/V_{beer}$	0.0455	0.01	0.060			
$N_{inf.sys.}/V_{beer}$	0.0455	0.01	0.060			
N_{sales}/V_{beer}	0.0455	0.01	0.060			
$N_{prod.}/V_{beer}$	0.0455	0.01	0.060			
$N_{econ.}/V_{beer}$	0.0455	0.01	0.060			
$N_{legal.prot.}/V_{beer}$	0.0455	0.01	0.060			
$N_{f.time}/V_{beer}$	0.0455	0.01	0.060			
$N_{p.time}/V_{beer}$	0.0455	0.01	0.060			
$N_{probationer}/V_{beer}$	0.0455	0.01	0.060			
N_{master}/V_{beer}	0.0455	0.581	0.058			
$N_{university}/V_{beer}$	0.0455	0.01	0.058			
$N_{college}/V_{beer}$	0.0455	0.01	0.058			
$N_{s.school}/V_{beer}$	0.0455	0.01	0.058			
$N_{skilled}/V_{beer}$	0.0455	0.01	0.058			
$N_{s.skilled}/V_{beer}$	0.0455	0.01	0.058			
$N_{un.skilled}/V_{beer}$	0.0455	0.01	0.058			
N_{total}/V_{beer}	0.0455	0.219	0.058			
Aspect: Occupational Health and Safety						
$N_{injuries}/V_{beer}$	0.0455	0.01	1.000			
Aspect: Diversity and Equal Opportunities						
$N_{total, male}/V_{beer}$	0.0455	0.01	0.250			
$N_{total, female}/V_{beer}$	0.0455	0.01	0.250			
$N_{age, male}/V_{beer}$	0.0455	0.01	0.250			
$N_{age, female}/V_{beer}$	0.0455	0.01	0.250			

Table 9 Weights of individual economic indicators and sub-index

	Indicators			Sub-index		
	EW	BOD	BAP	EW	BOD	BAP
Aspect: Economic Performance				0.333	0.1	0.348
R/V_{beer}	0.125	0.01	0.124			
$C_{operating}/V_{beer}$	0.125	0.93	0.130			
C_{salary}/V_{beer}	0.125	0.01	0.124			
$C_{dividend}/V_{beer}$	0.125	0.01	0.124			
C_{tax}/V_{beer}	0.125	0.01	0.124			
$C_{donation}/V_{beer}$	0.125	0.01	0.124			
$C_{package}/V_{beer}$	0.125	0.01	0.124			
$C_{env.tax}/V_{beer}$	0.125	0.01	0.124			

from the company, for example, the brewery did not have a probationer in 2007, (2) normalization of individual indicators using minimum–maximum method. According to this method the value of the normalized individual indicator was between 0 and 1. In the case of the indicator with a positive impact on sustainable development for

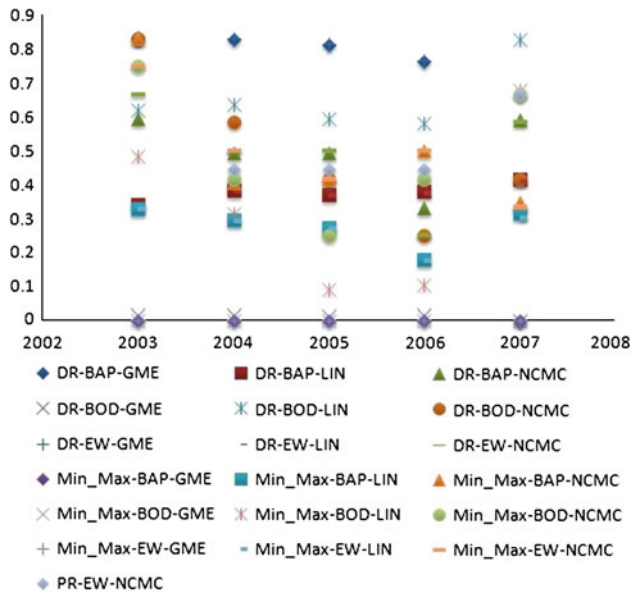


Fig. 3 Uncertainty analysis of different method combinations

$I_{i,j,t}^+ = I_{i,j}^{+,MIN}$, the normalized form had a value of 0 or in the case of the indicator had negative impact on sustainable development for $I_{i,j,t}^- = I_{i,j}^{-,MAX}$, according to the indicators' nature. Based on these observations it can be concluded that GME method should not be used if there is zero value in the origin data or in combination with minimum–maximum normalization method.

Sensitivity analysis can help to investigate which steps are the key drivers of the final results, the value of the composite sustainability index. During the calculation one of the method was fixed within a specific step within the construction frame and its impacts on the final results were examined, first-order sensitivity index. The first-order sensitivity index for the sustainability sub-indices and composite sustainability index is given in Fig. 4. According to the results, it can be concluded that when construction

sustainability sub-indices, different construction step will have the highest proportion in the output variance (the greatest impact on the result). In the case of the economic sub-index for the year 2003, the normalization step took part with more than 60% of the output uncertainty while in the case of the environmental sub-index for the same year (2003) the normalization step contributed to the output variance with only 26%. The aggregation method had the greatest impact on the environmental sub-index value for the year 2003. In general, the composite sustainability index was mainly affected by the normalization step (the first layer of data treatment), followed by the aggregation and weighting step.

The interactions among the construction steps can be measured by the total sensitivity index. In this analysis all methods were fixed except for one within the specific construction steps. The results are given in Fig. 5. The indices added up to a number greater than 1 due to the interactions which seemed to exist among the identified influential factors. As, $S_{Ti} > S_i$, it can be concluded that all of the construction combinations were non-additive. The difference $S_{Ti} - S_i$, points out that the aggregation step had more interaction with the other two construction steps, following by the normalization and the weighting step, Table 10.

Comparison between the two normalization methods, with reference to distance and minimum–maximum method is shown in Fig. 6, which suggests that the distance to the reference method had more influence on the output variance. The percentage of annual differences over the consecutive years method was excluded because of its limited compatibility and the information loss from the first year. Figure 7 gives a comparison among the weighting methods, EW, BAP, and BOD. According to the results, the 'benefit of the doubt' approach (BOD) gave the largest contribution to the output variance $V(Y)$. In the case of the aggregation methods, the GME was excluded from the

Fig. 4 The first-order sensitivity index for sustainability sub-indices and composite sustainability index

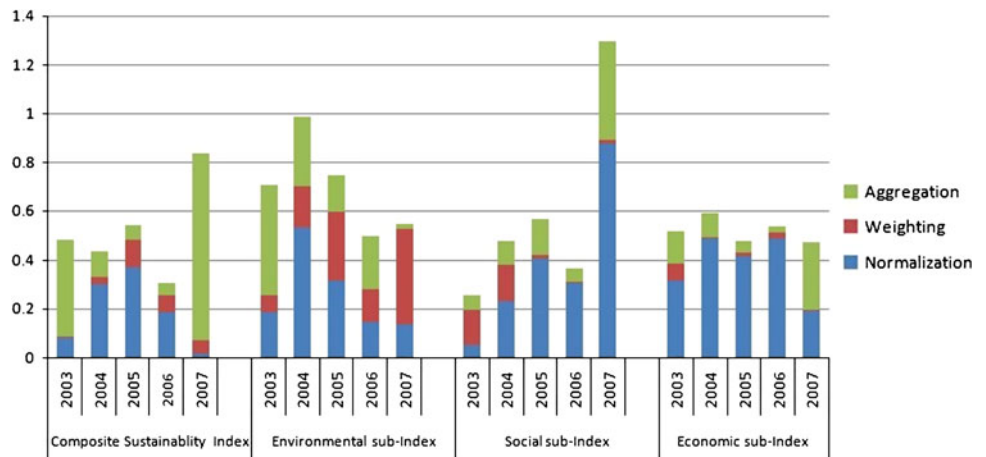


Fig. 5 The total sensitivity index of composite index

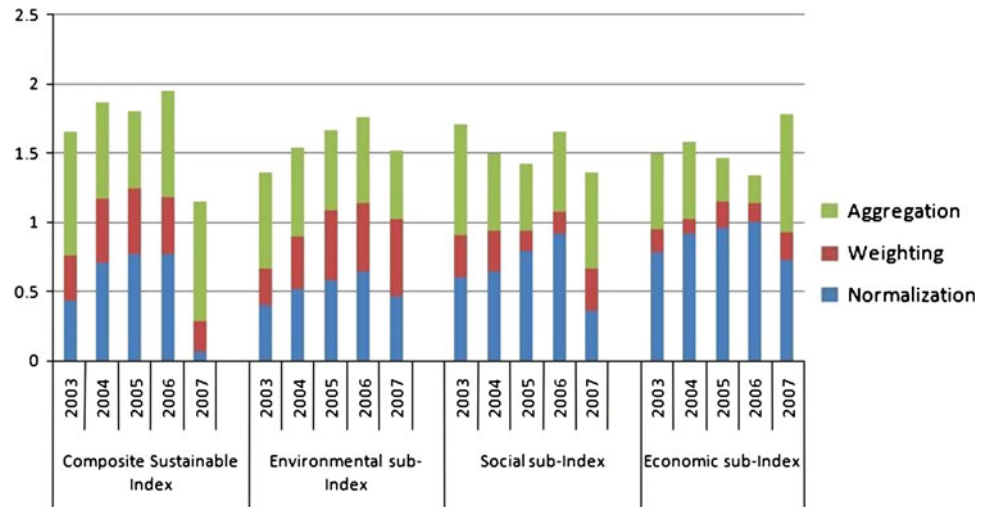


Table 10 First and total sensitivity index for the case study

	S_i	S_{Ti}	$S_{Ti} - S_i$
Normalization			
2003	0.0840	0.4345	0.3504
2004	0.3011	0.7085	0.4074
2005	0.3711	0.7670	0.3959
2006	0.1865	0.7669	0.5804
2007	0.0151	0.0645	0.0494
Weighting			
2003	0.0032	0.3220	0.3188
2004	0.0269	0.4584	0.4315
2005	0.1088	0.4735	0.3647
2006	0.0679	0.4144	0.3464
2007	0.0550	0.2228	0.1678
Aggregation			
2003	0.3931	0.8986	0.5055
2004	0.1086	0.7038	0.5952
2005	0.0609	0.5605	0.4996
2006	0.0505	0.7697	0.7192
2007	0.7660	0.8673	0.1013

investigation because of its disadvantage in dealing with zero involved data. A comparison of the remaining aggregation method is given in Fig. 8, which shows that LIN had the greatest effect on the final result.

Conclusions

This study investigated different combinations for calculating a composite sustainability index, and thus providing a comparison among different combinations to give a recommendation for the best way of constructing a composite index. Sustainable development indicators are

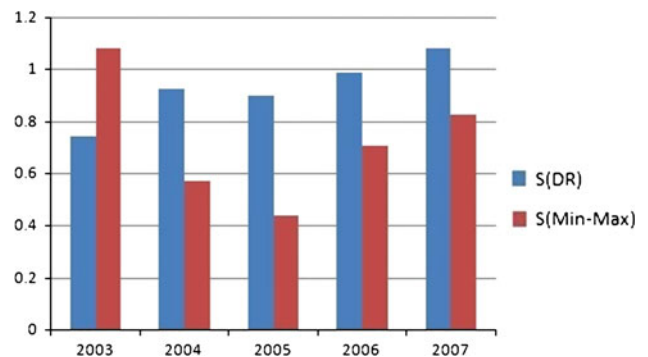


Fig. 6 Comparison of first-order sensitivity index among different normalization methods

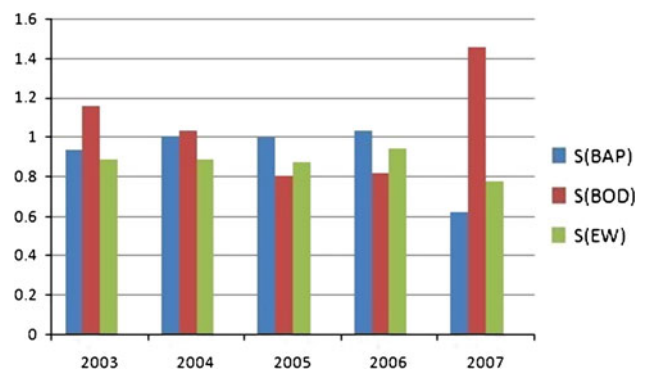


Fig. 7 Comparison of first-order sensitivity index among different weighting methods

recognized as useful tools for the assessment and anticipation of production performance and trends. Despite their increasing use, CIs remain controversial as they undesirably depend on the method used for normalizing the original data and causing disagreement among experts/stakeholders on the specific weighting scheme, and the way to aggregate sub-indicators. As a result, it is important to

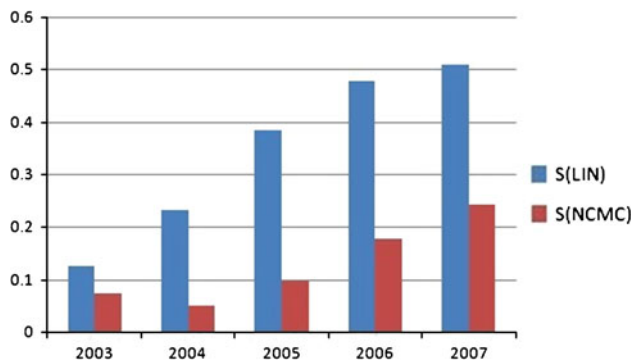


Fig. 8 Comparison of first-order sensitivity index among different aggregation methods

compare different combination of normalization–weighting–aggregation methods and create a recommendation for the most input sensitive construction scheme of CIs. The applied methodology in this article gradually aggregated sustainable development indicators into sustainability sub-indices and, finally, to a composite sustainability index. The normalization methods included in this analysis were: minimum–maximum, distance to a reference, and the percentage of annual differences over consecutive years. EWs, ‘benefit of the doubt’ approach, and BAP were used for determining the weights of individual indicators and sustainability sub-indices. The linear, geometric, and NCMCs were used in the cases of aggregation methods. Sensitivity analysis was performed over the results of 19 combinations of methods for constructing a composite sustainability index. According to the results, the normalization step had the largest influence on the composite sustainability index, followed by the aggregation and the weighting step. As the minimum–maximum normalization method will unavoidably result in a zero value causing information loss when combined with the GME method, this combination cannot be recommended for CI construction. For its limited compatibility and information loss during the first year, the percentage of annual differences over consecutive years is also not recommended. Within the normalization methods the ‘distance to a reference’ method had the greatest impact on the value of the composite sustainability index. In the weighting step, the ‘benefit of the doubt’ approach (BOD) and LIN methods for the aggregation step were identified as methods with the biggest influence on the result. Based on the conducted analysis and gained experience, the recommended construction frame is: the ‘distance to a reference’ normalization method in combination with the ‘benefit of the doubt’ and LIN.

Acknowledgment The authors are grateful to the National Natural Science Foundations of China (Project 21076180) for the financial support.

References

- Abusam A, Keesman K, Spanjers H, van Straten G (2004) Benchmarking procedure for full-scale activated sludge plants. *Control Eng Pract* 12(3):315–322
- Beloff BR, Tanzil D (2006) Assessing impacts: overview on sustainability indicators and metrics. *Environ Qual Manag* 15(4):41–56
- Chan K, Saltelli A, Tarantola S (1997) Sensitivity analysis of model output: variance-based methods make the difference. In: *Proceedings of the 1997 winter simulation conference*, pp 261–268
- Cherchye L (2007) An introduction to ‘benefit of the doubt’ composite indicators. *Soc Indic Res* 82(1):111–145
- Cherchye L, Moesen W, Rogge N, Van Puyenbroeck T (2006) Creating composite indicators with DEA and robustness analysis: the case of the technology achievement index. *J Oper Res Soc* 59(2):239–251
- De Carvalho SCP, Carden KJ, Armitage NP (2009) Application of a sustainability index for integrated urban water management in Southern African cities: case study comparison—Maputo and Hermanus. *Water SA* 35(2):144–151
- Frangopoulos CA, Keramioti DE (2010) Multi-criteria evaluation of energy system with sustainability considerations. *Entropy* 12(5):1006–1020
- Gasparatos A, El-Haram M, Horner M (2008) A critical review of reductionist approaches for assessing the progress towards sustainability. *Environ Impact Assess Rev* 28(4–5):286–311
- Geneletti D (2008) Impact assessment of proposed ski areas: a GIS approach integrating biological, physical and landscape indicators. *Environ Impact Assess Rev* 28(2–3):116–130
- Hatefi SM, Torabi SA (2010) A common weight MCDA-DEA approach to construct composite indicators. *Ecol Econ* 70(1):114–120
- Henning FPT, Muruvan S, Feng AW, Dunn CR (2011) The development of a benchmarking tool for monitoring progress towards sustainable transportation in New Zealand. *Transp Policy* 18(2):480–488
- Krajnc D, Mele M, Glavič P (2007) Fuzzy Logic Model for the performance benchmarking of sugar plants by considering best available techniques. *Resour Conserv Recycling* 52(2):314–330
- Mateos-Espejel E, Savulescu L, Maréchal F, Paris J (2011) Base case process development for energy efficiency improvement, application to a Kraft pulping mill. Part II: benchmarking analysis. *Chem Eng Res Des* 89(6):729–741
- Nardo M, Saisana M, Saltelli A, Tarantola S, Hoffman A, Giovannini E (2008) *Handbook on constructing composite indicators: methodology and user guide*. OECD Publishing, Paris
- Niemeijer D (2002) Developing indicators for environmental policy: data driven and theory driven approaches examined by example. *Environ Sci Policy* 5(2):91–103
- Phylipsen D, Blok K, Worrell E, de Beer J (2002) Benchmarking the energy efficiency of Dutch industry: an assessment of the expected effect on energy consumption and CO₂ emissions. *Energy Policy* 30(8):663–679
- Saisana M, Saltelli A (2008) Expert panel opinion and global sensitivity analysis for composite indicators. *Comput Sci Eng* 62:251–275
- Saltelli A (2002) Making best use of model evaluations to compute sensitivity indices. *Comput Phys Commun* 145(2):280–297
- Sikdar SK (2003) Sustainable development and sustainability metrics. *AIChE J* 49(8):1928–1932
- Sikdar SK (2009) On aggregating multiple indicator into a single metric for sustainability. *Clean Technol Environ Policy* 11(2):157–161
- Singh RK, Murty HR, Gupta SK, Dikshit AK (2007) Development of composite sustainability performance index for steel industry. *Ecol Ind* 7(3):565–588

- Singh RK, Murty HR, Gupta SK, Dikshit AK (2009) An overview of sustainability assessment methodologies. *Ecol Ind* 9(2):189–212
- Tokos H, Novak Pintarič Z, Krajnc D (2011) An integrated sustainability performance assessment and benchmarking of breweries. *Clean Technol Environ Policy*. doi:[10.1007/s10098-011-0390-0](https://doi.org/10.1007/s10098-011-0390-0)
- von Bahr B, Hanssen OJ, Vold M, Pott G, Stoltenberg-Hansson E, Steen B (2003) Experiences of environmental performance evaluation in the cement industry. Data quality of environmental performance indicators as a limiting factor for Benchmarking and Rating. *J Clean Prod* 11(7):713–725
- Zhou P, Ang BW, Poh KL (2007) A mathematical programming approach to constructing composite indicators. *Ecol Econ* 62(2):291–297
- Zhou P, Ang BW, Zhou D (2010) Weighting and aggregation in composite indicator construction: a multiplicative optimization approach. *Soc Indic Res* 96(1):169–181