METHODS

A mathematical programming approach to constructing composite indicators

P. Zhou*, B.W. Ang, K.L. Poh

Department of Industrial and Systems Engineering, National University of Singapore, 10 Kent Ridge Crescent, 119260, Singapore

ARTICLE INFO

Article history:
Received 5 September 2006
Received in revised form
7 November 2006
Accepted 22 December 2006
Available online 15 February 2007

Keywords:
Composite indicators
Aggregation
Weighting
Data envelopment analysis
Sustainable energy development

ABSTRACT

Composite indicators (CIs) have been widely accepted as a tool for performance monitoring, benchmarking, policy analysis and public communication in various fields. To a large extent, the usefulness of a CI depends heavily on the underlying weighting and aggregation schemes. In this paper, we propose a mathematical programming approach to constructing composite indicators. The proposed approach uses two sets of weights that are most and least favourable for each entity to be evaluated and therefore could provide a more reasonable and encompassing CI. We apply the proposed approach to develop a CI for modeling the sustainable energy development of eighteen APEC economies and present the results obtained.

© 2007 Elsevier B.V. All rights reserved.

1. Introduction

A composite indicator (CI) is a mathematical aggregation of a set of individual indicators that measure multi-dimensional concepts but usually have no common units of measurement (Nardo et al., 2005). Despite the ceaseless debate on its use, CI has been increasingly used for performance monitoring, benchmarking, policy analysis and public communication in wide ranging fields including economy, environment and society by many national and international organizations. Examples of well-known CIs include the UN’s Human Development Index (Sagar and Najam, 1998), and the Environmental Performance Index produced by a joint effort among Yale, Columbia, World Economic Forum and the Joint Research Center of European Commission (Esty et al., 2006). More discussions on various CIs can be found in the information server: http://farmweb.jrc.cec.eu.int/ci/ provided by the Joint Research Center of European Commission.

The popularity of CIs in practice is likely due to what has been pointed out by Saisana et al. (2005): the temptation of stakeholders and practitioners to summarize complex and sometime elusive process (e.g. sustainability or a single-market policy) into a single figure to benchmark country performance for policy consumption seems irresistible. Nevertheless, to a large extent, the usefulness of a CI depends heavily on the underlying weighting and aggregation schemes. Therefore, the study on data weighting and aggregation has always been an interesting but controversial topic in the field of CI construction (Esty et al., 2006).

At the stage of data aggregation, the applicability of multiple criteria decision analysis (MCDA) methods has been widely investigated. See, for example, Munda and Nardo (2003), Díaz-Balteiro and Romero (2004), Ebert and Welsch (2004), Krajnc and Glaviè (2005), Munda (2005), Nardo et al. (2005), Hajkowicz (2006), Zhou et al. (2006a). A major problem in applying MCDA aggregation methods to construct a CI is the determination of the weights for the underlying sub-
indicators. From the methodological viewpoint, there exist many weighting methods which can be used to derive the weights for sub-indicators. Nardo et al. (2005) have recently discussed the pros and cons of different weighting methods. In practice, expert judgment or public opinion poll results are often used to derive the weights for sub-indicators (Hope et al., 1992). When such information is unavailable, as illustrated in the Human Development Index (1992). When such information is unavailable, as illustrated in the Human Development Index (1992), Lau and Lam (2002), Cherchye et al. (in press-a,b), Lovell et al. (1995), Mahlberg and Obersteiner (2001), Cherchye (2001), Lau and Lam (2002), Cherchye et al. (in press-a,b), Despotis (2005a,b). Our paper also follows this line of research. More specifically, this paper extends previous studies and presents a simple mathematical programming approach to constructing CIs in virtue of the idea of DEA.

2. Problem description

Consider the case where there are \( m \) entities, e.g. countries or regions, whose aggregated performance are to be evaluated based on \( n \) sub-indicators. These sub-indicators usually have no common measurable units. Let \( I_i \) denote the value of entity \( i \) with respect to sub-indicator \( j \). Without loss of generality, we further assume that all the sub-indicators are of the benefit type which satisfy the property of “the larger the better”. As illustrated in Fig. 1, the problem is to aggregate \( I_i, (i=1,2,...,n) \) into a composite index \( CI \), that can be used to evaluate the aggregated performance of entity \( i \) with respect to all the underlying sub-indicators.

A common practice in constructing CI is to first assign a weight to each sub-indicator, and then use certain aggregation functions to calculate CIs. Before aggregation the normalization of sub-indicators is often required. Two simple but popular aggregation methods that have often been adopted are the simple additive weighting (SAW) method and the weighted product (WP) method (Nardo et al., 2005). Munda and Nardo (2003) have recently pointed out that non-compensatory MCDA approach may have some theoretical advantages over other aggregation methods in CI construction. In view of the fact that there exist many aggregation methods, Ebert and Welsch (2004) and Zhou et al. (2006a) have provided some guidelines for aggregation method selection. Nevertheless, the assignment of weights is still a controversial issue in CI construction for many practitioners and researchers.

3. The proposed approach

As has been mentioned earlier, a critical issue in using MCDA aggregation methods to construct CIs is the subjective in assigning weights for sub-indicators. Since different weight combinations may lead to different ranking results, it is unlikely that all the entities would easily reach a consensus in determining an appropriate set of weights. In addition, it may not be easy to obtain the expert information for deriving

\[
\begin{bmatrix}
I_{11} & I_{12} & \cdots & I_{1n} \\
I_{21} & I_{22} & \cdots & I_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
I_{m1} & I_{m2} & \cdots & I_{mn}
\end{bmatrix} \rightarrow \begin{bmatrix}
CI_1 \\
CI_2 \\
\vdots \\
CI_m
\end{bmatrix}
\]

Fig. 1 – Graphical representation of CI construction.
the weights. Although the use of equal weights seems to be a relatively fair choice, some entities may still have different opinions since they have their particular preferences. To avoid these issues, a DEA-like model is given here for aggregation purpose:

\[
g_i = \max_{j=1}^{n} w_{ij} g_{ij}
\]

subject to

\[
\sum_{j=1}^{n} w_{ij} \leq 1, \quad k = 1, 2, \ldots, m
\]

\[
w_{ij} \geq 0, \quad j = 1, 2, \ldots, n
\]

(1)

Model (1) provides an aggregated performance score for entity \(i\) in terms of all the underlying sub-indicators. By solving Model (1) repeatedly for each entity, we will obtain a set of indices \(g_i, g_{i2}, \ldots, g_{in}\) for these entities. Note that the objective function in Model (1) is externally similar to the objective function in Model (2) and the weights for sub-indicators are endogenous and changeable while in the SAW method they are exogenous and fixed. In essence, Model (1) is an output maximizing multiplier DEA model with multiple outputs and constant inputs, which measures how far the evaluated entity is from the best practice entity under the best possible weights. It should be pointed out that Model (1) is not a new model, i.e. Models (1) and (2) are similar in approach to reassessing the "best" set of weights for each entity which are used to aggregate the sub-indicators into a performance score. Externally, Model (2) measures how close the evaluated entity is from the worst practice entity under the worst possible weights. It provides a way for further performance comparison among those incomparable entities only based on Model (1). It is worth pointing out that Model (2) is not a brand-new model in the DEA literature. Conceptually, it is parallel to the minimum efficiency concept as discussed in Zhu (2004). The similar idea has also been applied to study a site selection problem by Takamura and Tone (2003). Nevertheless, as far as we know, it is the first time that Model (2) is applied to the field of CI construction.

So far we have provided two performance indexes for each entity which are derived from the data by two DEA-like models, i.e. Models (1) and (2). Since the two indexes are based on weights that are most favourable and least favourable for each entity, they could only reflect partial aspects of an entity in terms of its aggregated performance. It is logical and reasonable to combine them into an overall index. Therefore, we combine the two indexes to form a CI in the following way:

\[
CI(\lambda) = \lambda \frac{gl_i-gl^*}{gl^*-gl_i} + (1-\lambda) \frac{bl_i-bl^*}{bl^*-bl_i}
\]

where \(gl^* = \max\{g_i, i=1,2,\ldots, m\}\), \(gl^* = \min\{g_i, i=1,2,\ldots, m\}\), \(bl^* = \max\{b_i, i=1,2,\ldots, m\}\), \(bl^* = \min\{b_i, i=1,2,\ldots, m\}\), and \(0 \leq \lambda \leq 1\) is an adjusting parameter.

In Model (3), we use the linear scaling in the min–max range to let the two indexes become comparable (in the range of [0, 1]) and then use the linear aggregation to combine them together by the adjusting parameter \(\lambda\). If \(\lambda = 1\), CI will become a normalized version of \(gl_i\). If \(\lambda = 0\), CI will become a normalized version of \(bl_i\). For other cases, Model (3) makes a compromise between the two indexes. If decision makers or analysts have no particular preference, \(\lambda = 0.5\) seems to be a fairly neutral choice. Given these characteristics of Model (3), we may believe that Model (3) provides a more encompassing CI since it takes into account two extreme cases. Nevertheless, the indexes given by Models (1) and (2) should not be discarded since they can provide some other valuable information such as the "performance bound" information.

It can be easily shown that CI satisfies a number of desirable properties:

P1. 0 ≤ CI ≤ 1.

P2. CI is units invariant.

---

2 It implies that in our approach we adopt linear rather than nonlinear aggregation procedure. Although nonlinear aggregation may have some advantages as discussed in Ebert and Welsch (2004) and Zhou et al. (2006a), we have chosen linear aggregation procedure because of its simplicity and ease of understanding. In addition, linear aggregation is a common practice in DEA literature.

3 It should be pointed out that linear scaling in the min–max range and the linear aggregation are not the only scaling and aggregation approaches available. However, considering the fact that linear aggregation is a common practice in DEA and it is often based on the normalized values from linear scaling in the min–max range in MCDA, we have chosen them in favor of other approaches.
P3. CI is invariant to the right hand sides of the constraints in Models (1) and (2).

The property P1 indicates that Model (3) provides a standardized index which lies in the interval [0, 1]. If an entity has a larger value, we may say that this entity has better aggregated performance. If an entity has the largest values in terms of both $g_i$ and $b_i$, it will give a CI of “1” no matter what $\lambda$ is. If an entity has the smallest values in terms of both $g_i$ and $b_i$, it will give a CI of “0” no matter what $\lambda$ is. The property P2 says that whatever units are chosen for sub-indicators, the value of CI will remain unchanged. As a result, we do not need to consider the normalization procedure before aggregation in applying the proposed approach to construct CIs. P3 implies that if we replace “1” in the constraints of Models (1) and (2) by any other positive values, the value of CI would remain unchanged.

Note that in the previous models no exogenous restrictions are imposed on the weights for sub-indicators. All the weights are generated from the data. Under these circumstances the weights used may be such that a number of sub-indicators would be ignored in aggregation. This is definitely not the case we expect since all the sub-indicators selected should be considered as “important” ones and it may not be appropriate to ignore many of them. To overcome this problem, we may consider restricting the flexibility of weights in an appropriate way by incorporating additional information. In principle, this can be done by a number of methods as reviewed by Allen et al. (1997). Mahlberg and Obersteiner (2001), Cherchye et al. (in press-a,b) have given a few demonstrations on how to restrict the flexibility of weights in CI construction.

Although a number of studies highlight the direct restrictions on weights, we would suggest the use of “proportion constraints” proposed by Wong and Beasley (1990) in the DEA literature. Technically, we can revise Models (1) and (2) by respectively adding the following two sets of constraints:

$$L_j \leq \sum_{i=1}^{n} \frac{u_{ij}^I I_{ij}}{v_{ij}^I} \leq U_j, \quad j = 1, 2, \ldots, n$$

$$L_j \leq \sum_{i=1}^{n} \frac{u_{ij}^I I_{ij}}{v_{ij}^I} \leq U_j, \quad j = 1, 2, \ldots, n$$

where $L_j$ and $U_j$ are respectively denote the lower and upper limits for the contribution of the $j$-th sub-indicator in CI and satisfy $0 \leq L_j \leq U_j \leq 1$. The main reason for using this way to restrict the flexibility of weights arises from some practical considerations. As Cherchye et al. (in press-a,b) argued, it is easier and more practical to let experts make a “limited agreement” on the determination of weights. Therefore, it is not a difficult task to derive the limits $L_j$ and $U_j$ in practice. Usually this can be done by making a consensus among decision makers or domain experts as to the relative importance of each sub-indicator. In the case that no consensus could be reached in terms of a certain sub-indicator, we can remove the corresponding weight restriction constraint. If no expert information is given, we can let $L_j=0$, $U_j=1 (j=1,2, \ldots, n)$ and the revised models will reduce to Models (1) and (2). Another advantage of using “proportion constraints” is that it preserves the desirable units’ invariance property as pointed out by Cherchye et al. (in press-a,b).

### 4. Case study: sustainable energy index

Sustainable energy development, which consists of such elements as energy supply, energy efficiency and environmental protection, is a major concept in sustainable...
development (Jefferson, 2006). It is therefore very meaningful to gauge the sustainable energy development of a country/region relative to other countries/regions. In this section, we shall apply the proposed approach to develop a CI, namely sustainable energy index (SEI) for eighteen APEC economies in 2002 for measuring and comparing their performance towards sustainable energy development. Using this application study, we can also illustrate what the general procedure for constructing CIs is and how the proposed approach can be used to construct CIs in practice.

The first step for constructing SEI is to select appropriate underlying sub-indicators. The International Atomic Energy Agency (IAEA) has recently published a total of thirty indicators related to sustainable energy development (IAEA, 2005). Despite its comprehensiveness, to include all of them is impossible due to the lack of data. Following Esty et al. (2006), we choose only energy efficiency indicator (EEI), renewable energy indicator (REI) and climate change indicator (CCI) as our sub-indicators for constructing SEI.

In the case of energy efficiency, various indicators, e.g. thermodynamic indicators, physical-based indicators and monetary-based indicators, have been used. According to Ang (2006), monetary-based indicators are more suitable for measuring energy efficiency at a high level of aggregation, which corresponds to our case here. We therefore choose the energy–GDP ratio, e.g. the ratio of total final energy consumption to GDP, as our EEI. The data on energy consumption and GDP are collected from APEC Energy Statistics 2003 (APEC, 2005). The REI is defined as the percentage of renewable energy in total final energy consumption. Renewable energy consumption includes such renewable sources as hydro-electricity, geothermal, solar and wind, and the data are collected from the US Energy Information Administration (2005). In the case of CCI, we follow Esty et al. (2006) and use the ratio of CO2 emissions to GDP as its proxy. The data on CO2 emissions are collected from the World Resources Institute (2005).

Note that according to our previous definitions EEI and CCI are cost-type indicators. So we first transform them into benefit-type indicators by taking their reciprocals before aggregation. We then apply Models (1), (2) and (3) to calculate the SEI values of the eighteen economies. Table 1 presents the results obtained as well as the data for the three sub-indicators.

It can be seen from Table 1 that all the economies can be compared with each other based on the SEI values while four of them have the same performance score of “1” by Model (1), which indicates that the proposed approach could lead to CIs with higher discriminating power. Ignoring the issue of data quality, Table 1 could also provide some useful information about sustainable energy development in APEC economies. It is observed from Table 1 that Peru has the highest SEI value (=1) although none of the three sub-indicators rank first in the country. It is likely due to the fact that Peru not only has relatively high sub-indicator values but also has a better balance among different sub-indicators. From Table 1 we can also observe that Russia has the least SEI value (=0) since Russia’s two indexes given by Models (1) and (2) are the smallest compared to other economies. As a whole, it seems that the sustainable energy development of APEC economies is not so good since most economies have a very small SEI value and the average is below 0.4.

Note that in Table 1 we calculate the SEI values by fixing the adjusting parameter $\lambda$ at 0.5. To investigate whether the adjusting parameter has severe effects on SEI, we further consider the cases that $\lambda=0.1, 0.2, ..., 0.9$. Using the nine $\lambda$ we can get nine SEI scores for each of the eighteen economies.

---

**Fig. 2** – Comparative box plots of SEI values for eighteen APEC economies in 2002.

**Fig. 3** – Comparative box plots of SEI ranks for eighteen APEC economies in 2002.

**Fig. 4** – Comparison between the SEI by basic models and that by models with weight restrictions.
Table 2 – Correlation between sub-indicators and SEI

<table>
<thead>
<tr>
<th>Scenario 1</th>
<th>EEI</th>
<th>REI</th>
<th>CCI</th>
<th>SEI</th>
<th>Scenario 2</th>
<th>EEI</th>
<th>REI</th>
<th>CCI</th>
<th>SEI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model (1)</td>
<td>0.660 (0.636)</td>
<td>0.876 (0.792)</td>
<td>0.832 (0.872)</td>
<td>1 (1)</td>
<td>Model (1)</td>
<td>0.836 (0.680)</td>
<td>0.803 (0.833)</td>
<td>0.911 (0.868)</td>
<td>0.929 (0.981)</td>
</tr>
<tr>
<td>Model (2)</td>
<td>0.800 (0.670)</td>
<td>0.841 (0.808)</td>
<td>0.910 (0.878)</td>
<td>0.961 (0.992)</td>
<td>1 (1)</td>
<td>Model (2)</td>
<td>0.836 (0.680)</td>
<td>0.803 (0.833)</td>
<td>0.911 (0.868)</td>
</tr>
<tr>
<td>Model (3)</td>
<td>0.852 (0.701)</td>
<td>0.794 (0.827)</td>
<td>0.918 (0.872)</td>
<td>0.910 (0.979)</td>
<td>0.995 (0.996)</td>
<td>1 (1)</td>
<td>Model (3)</td>
<td>0.800 (0.670)</td>
<td>0.841 (0.808)</td>
</tr>
</tbody>
</table>

Note: The Spearman rank correlation coefficient is included in parentheses.

Fig. 2 shows the comparative box plots of SEI values for the eighteen economies in the sequence of the mean SEI values, and Fig. 3 shows the comparative box plots of their SEI ranks in the same sequence. It can be seen from Fig. 2 that the SEI value is very insensitive to λ for most economies. In fact, as shown in Fig. 3, the ranks of over two thirds of economies in terms of their sustainable energy development are fairly consistent under different λ values. The Spearman ranking correlation coefficients between the SEI values (λ=0.5) and the SEI values in other cases are all larger than 0.98.

Previous discussions are based on the basic models of the proposed approach, i.e. Models (1), (2) and (3). In the following, we shall consider the case that the flexibility of weights is restricted in the form of Models (4) and (5). Since the current case study is mainly for illustration purposes, we arbitrarily choose \( L_1=L_2=L_3=0.1 \) and \( U_1=U_2=U_3=0.5 \) for use, which indicates that the contribution of each sub-indicator is not less than 10% but not larger than a half of the aggregated CI. We then apply the resulting models to recalculate the SEI values for these economies by using \( \lambda=0.5 \). The results obtained, labeled as Scenario 2, as well as the SEI values without restricting the flexibility of weights (Scenario 1) are displayed in Fig. 4.

It can be seen from Fig. 4 that the SEI values of many economies, e.g. Peru, the Philippines, Thailand and China have no or little changes after weights for sub-indicators are restricted. On the contrary, in some economies there have been obvious changes. For instance, under Scenario 1 Russia has the least SEI value while under Scenario 2 Korea has the least SEI value. This could be explained by the fact that Korea has a very small REI value and the weight for REI is zero by Model (1) under Scenario 1, while under Scenario 2 it is larger than zero.

Table 2 shows the correlations of the two sets of SEI values under different scenarios with each other and with the three sub-indicators. The correlations of the two sets of indexes given by Model (1) with each other, with the two sets of SEI values, and with the three sub-indicators are also displayed in Table 2. It can be observed from Table 2 that the contributions of various sub-indicators in SEI are obviously different. It seems that SEI is strongly connected with CCI in both scenarios. When the weight restrictions are considered, the indexes given by Model (1) become more highly correlated with EEI, CCI and the SEI values under Scenario 1. It is likely due to the fact that the discriminating power of Model (1) becomes higher when considering weights restrictions. In the case of rank correlation, we can find that the SEI ranks under Scenario 2 are more consistent with the ranks of sub-indicators. We can also find that the two sets of SEI values are highly correlated with each other, which may be an indication of the robustness of the proposed approach in constructing CIs.

5. Conclusion

CIs have been widely accepted as a useful tool for performance monitoring, benchmarking, policy analysis and public communication in wide ranging fields, e.g. economy, environment and society. In this paper, we propose a mathematical programming approach to constructing CIs. Compared with previous studies, the proposed approach requires no prior knowledge of the weights for sub-indicators. The weights used can be generated by solving a series of DEA-like models. Since the proposed approach uses two sets of weights that are most and least favourable for each entity, it may provide a more reasonable and encompassing CI. In addition, the proposed approach can easily incorporate additional information on the relative importance of sub-indicators when they are available.

The proposed approach has been applied to develop a CI for modeling the sustainable energy development of eighteen APEC economies in 2002. We first apply the basic models of the proposed approach to construct a SEI for each economy. We then investigate whether the adjusting parameter used in the proposed approach has severe effects on the SEI values. It is found that the SEI value is very insensitive to this parameter. The scenario in which the flexibility of weights is restricted has also been investigated and the results obtained are compared with the scenario in which basic models are used. It is found that the two sets of SEI values are highly correlated with each other, which may be an indication of the robustness of the proposed approach in constructing CIs.

Acknowledgement

The authors would like to thank two anonymous referees for their valuable comments on the earlier version of this paper. The usual disclaimer applies.

REFERENCES


