Journal of the Operational Research Society (2013), 1–12

www.palgrave-journals.com/jors/

Composite indicator development using utility function and fuzzy theory

S-K Lee and J-H Yu*

Kwangwoon University, Seoul, South Korea

Construction companies use composite indicators (CIs) to evaluate their overall project performance. However, the conventional methodology of CIs development causes indiscrimination, relative calibration, and redundancy. To address these problems, we propose a novel methodology that uses fuzzy theories. The proposed methodology includes a utility function for normalizing, a fuzzy measure for weighting, and a fuzzy integral for aggregating. We conducted a case study to assess the quality of the proposed methodology *versus* the alternative methodologies on 25 real projects of a construction company. The result showed that the measurement reliability of the proposed normalization method (1.96) is greater than that of the two different normalization methods (10.44 and 2.8, respectively). In addition, the measurement accuracy of the proposed aggregation method is greater than those of the four different aggregation methods. Therefore, our proposed methodology can more consistently and accurately help evaluate the overall project performance or success.

Journal of the Operational Research Society advance online publication, 20 February 2013
doi:10.1057/jors.2013.15

Keywords: multi-criteria; composite indicator; utility function; fuzzy integral; fuzzy measure uncertainty analysis

1. Introduction

Construction companies evaluate project success by measuring performance and comparing it with that of other projects according to predetermined success criteria. These criteria include schedule, cost, quality, and safety performance; each aspect has many sub-indicators to measure its performance (Kumaraswamy and Thorpe, 1996; Dainty *et al*, 2003). To evaluate the overall project performance or success, construction companies have developed composite indicators (CIs), in which sub-indicators are aggregated into one index. Construction companies commonly use a categorical scale, *Z*-score, or re-scaling to normalize the values of sub-indicators with different measures; a budget allocation to weight the sub-indicators; and a simple additive aggregation function to aggregate the weighted sub-indicators.

However, despite their simplicity in implementation and interpretation, these methods do not appropriately address their inherent problems. For instance, the categorical scale converts the continuous values of the sub-indicators into discontinuous categorical values, and the low resolution measurement often impairs the performance discrimination in the process (Hand, 2004). Although the normalized values by Z-score or re-scaling are continuous, these methods provide relative calibration due to their nature. Therefore, the values obtained by these methods differ according to project performance. Moreover, the simple additive weighting method does not consider that the interaction among sub-indicators can cause redundancy (Grabisch, 1996). Developing a CI by merely adding the weights of these indicators can lead to an incorrect estimation of safety performance due to the redundancy in these two sub-indicators.

We address these problems by developing a novel methodology that applies fuzzy theories to develop a CI for evaluating overall project performance. Specifically, we propose the utility function as a replacement for the categorical scale, *Z*-score, or re-scaling to address the indiscrimination and relative calibration. We then apply the fuzzy integral to aggregate the normalized sub-indicators in order to avoid redundancy. We assess the reliability and accuracy of the proposed approach using uncertainty analysis. The test case is the cost performance of 25 real projects provided by a construction company in Korea.

The rest of the paper is organized as follows. Section 2 discusses the CI of the construction project performance and the methodologies for constructing the CI. We then propose a fuzzy-based methodology for the CI in Section 3. In Section 4, we demonstrate the quality of our proposed

Correspondence: J-H Yu, Department of Architecture Engineering, School of Engineering, Kwangwoon University, 447-1, Wolgye-Dong, Seoul, Nowongu 139-701, South Korea.
E-mail: myazure@kw.ac.kr

methodology using a Monte Carlo approach-based uncertainty analysis. Finally, we conclude and offer some final remarks in Section 5.

2. Review of constructing CIs

2.1. Current practice of constructing CIs

Construction companies have utilized CIs to measure and compare their overall project performance due to their usefulness as a communication tool (Freudenberg, 2003) and as a decision support tool (Saltelli, 2006). Because only overall project performance can be measured using a CI, many researchers support the use of both composite performance indicators and individual indicators (ie, project success criteria or key performance indicators). Lauras *et al* (2010) and Marques *et al* (2010) argue that project managers need to quantify project performance as a whole. Clivillé *et al* (2007) state that the performance management system should involve two kinds of performance metrics: elementary (ie, individual indicators that represent different performance objectives) and aggregated (ie, CIs that synthesize the elementary indicators into global objectives). Kumaraswamy and Thorpe (1996) suggest the use of a project performance profile composed of principal performance criteria and corresponding sub-criteria in a hierarchical structure. Landy and Farr (1983) argue that combined performance data are needed, because the availability of overall performance ratings is useful for administrative decisions.

To develop a CI for evaluating overall project performance, construction companies often use a categorical scale for normalization, a budget allocation for weighting, and a simple additive aggregation function for aggregation. Although these methods are widely used in the development of a CI (Saisana and Tarantola, 2002), they assume preference independence, which Nardo *et al* (2005) define as 'given the sub-indicators, an simple additive aggregation function exists if and only if these indicators are mutually preferentially independent'. If two or more indicators measure the same system behaviour or violate the preference independence assumption, a certain performance aspect will be redundantly weighted (Grabisch, 1996; Freudenberg, 2003). To address this redundancy, interrelations between the sub-indicators must be taken into account when the sub-indicators are weighted and aggregated.

2.2. Methodology of CI construction

A CI is generally developed by developing a theoretical framework, selecting sub-indicators, inventing a CI, testing the robustness of the CI, and using the CIs to report the results (Saisana and Tarantola, 2002; Freudenberg, 2003; Nardo *et al*, 2005; OECD and European Commission-Joint Research Center, 2008). Inventing a CI consists of three steps: normalization, weighting, and aggregation. Various methods have been developed for each step: (a) normalization methods: a number of normalization methods exist (Freudenberg, 2003; Jacobs and Goddard, 2004). Standardization (or Z-scores) converts indicators to a common scale with a mean of zero and a standard deviation of one. Thus indicators with extreme values have a greater effect on the CI. Re-scaling normalizes indicators to have an identical range [0, 1] by subtracting the minimum value and dividing it by the range of the indicator values. However, extreme values/or outliers could distort the transformed indicator. (b) Weighing methods: weighting schemes range from statistical models (such as factor analysis, data envelopment analysis, and unobserved component models) to participatory methods (such as budget allocation or analytic hierarchy processes). Weights usually have an important impact on the composite value and on the resulting ranking especially whenever higher weight is assigned to the sub-indicators. (c) Aggregation methods: the simple additive weighting (SAW) method, the weighted product (WP) method, the weighted displaced ideal (WDI) method, and the technique for order preference by similarity to ideal solution (TOPSIS) method have also been widely explored in CI construction (Diaz-Balteiro and Romero, 2004; Ebert and Welsch, 2004; Esty *et al*, 2005; Nardo *et al*, 2005a, b; Zhou *et al*, 2006; Lun *et al*, 2006). Although the TOPSIS method has been rarely used to construct CIs, it has attractive properties (Yoon and Hwang, 1995; Sinha and Shah, 2003). These methods provide the opportunity to choose an appropriate set of methods based on the context of the evaluation. Researchers (Park *et al*, 2009; Shouke *et al*, 2010; Bai *et al*, 2011; Cha and Kim, 2011) suggest various CI models, which are different from the widely accepted model in the construction industry. However, set of methods for addressing indiscrimination, relative calibration, and redundancy problems remains lacking.

Fuzzy theories, including the fuzzy measure and the fuzzy integral, can be utilized to address these problems due to their ability to model the interaction among sub-indicators (Grabisch, 1996). The Choquet and the Sugeno integrals are two well-known forms of the fuzzy integral. While the Sugeno integral is based on nonlinear operators (min and max), the Choquet integral is based on linear operators and is a natural extension of the Lebesgue integral (Liginlal and Ow, 2006). Many researchers apply fuzzy theories in various disciplines such as evaluating enterprise intranet websites (Tzeng *et al*, 2005) and e-commerce strategies (Chiu *et al*, 2004). In the construction industry, fuzzy theories have been used to manage uncertainties in design performance prediction (Fayek and Sun, 2001) or labour productivity (Fayek and Oduba, 2005). Although these studies provide valuable insight into the relationships between fuzzy theories and performance evaluation, they do not explicitly address the indiscrimination and redundancy problems in the context of construction project performance evaluation. Research efforts that apply fuzzy theories to evaluate overall project performance and explain application effectiveness are needed.

3. Fuzzy-based methodology for CI

There is a need for a novel methodology to help construction companies develop a CI that addresses the indiscrimination, relative calibration, and redundancy problems by applying fuzzy theories in synthesizing multiple criteria.

The proposed methodology has the following steps:

Step 1 (*Normalization*): To address the problem of indiscrimination and relative calibration during the process of normalizing the values of the sub-indicators, a utility function is used as a normalization method to combine the values into a composite value. The utility value is measured in arbitrary units called utiles. The x-axis (the utility function's argument) is calibrated in directly measureable units. The y-axis origin and scale (expressed in utiles or utils) are arbitrary (Schuyler, 1996). The utility function can help address the indiscrimination problem, because the y-axis can also have continuous values. For the y-axis, a 0–1 scale can be used to normalize different scales of sub-indicators without affecting the discriminating power of these sub-indicators. This function interpolates the values within a given category using two boundary conditions that represent a company's perception of the utility. Although the use of utility functions that represent a construction company's preference would produce more realistic normalization results, we use the 0–1 scale for the utility functions for demonstrative purposes.

Step 2 (*Weighting*): The normalized values are weighted using the fuzzy measure. The method used to obtain λ-fuzzy measure values for the Choquet fuzzy integral is as follows. First, we determine g_i which is the importance measure or the contribution of each single sub-indicator to a CI. The fuzzy measure can be used to model the interrelation between sub-indicators. Therefore, there is no need to include the constraint that the sum of influence of each sub-indicator must be one. Second, we calculate the value of λ using Equation (1) given the g_i determined above.

$$1+\lambda = \Pi_{i=1}^{n}(1+\lambda g_i), \quad \lambda \neq 0, -1<\lambda \tag{1}$$

In addition, according to the fundamental theorem regarding the λ-fuzzy measure, λ-value has the following cases:

- If $\sum_{i=1}^{n} g_i > g(X)$, then $-1<\lambda<0$
- If $\sum_{i=1}^{n} g_i = g(X)$, then $\lambda=0$
- If $\sum_{i=1}^{n} g_i < g(X)$, then $\lambda>0$

Third, the values of normalized sub-indicator $h(x_i)$ are listed in descending order, and we calculate the λ-fuzzy measure value of each $g(H_i)$ using the λ, g_i values and Equation (2).

$$g(H_i) = g(\{x_i, x_{i+1}, \ldots, x_n\}) = \frac{1}{\lambda}\left[\Pi_{j=i}^{n}(1+\lambda g_j) - 1\right] \tag{2}$$

where $g_i = g(\{x_i\})$, $g_j = g(\{x_j\})$, $H_i = \{x_i, x_{i+1}, \ldots, x_n\}$, and $i = 1,2, \ldots, n$

Step 3 (*Aggregation*): We suggest the use of the Choquet integral for aggregating the sub-indicators in our proposed methodology. The Choquet fuzzy integral, proposed by Murofushi and Sugeno (1989), has been used in information fusion and data mining as a nonlinear aggregation tool (Yang *et al*, 2005). This method provides the computational schemes for aggregating the values of sub-indicators based on the λ-fuzzy measure described above. If $h(x_1)$, $h(x_2), \ldots, h(x_n)$ are assumed to be a collection of input sources of h, and g is a λ-fuzzy measure, then the following Choquet fuzzy integral can be constructed:

$$\int_x h(x)^{\circ} g(\cdot) = \Sigma_{i=1}^{n}[h(x_i) - h(x_{i-1})]g(H_i) \tag{3}$$

where x is a finite and discrete set, $H_i = \{x_i, x_{i+1}, \ldots, x_n\}$, $h(x_1) \leqslant h(x_2) \leqslant \ldots \leqslant h(x_n)$ and $h(x_0) = 0$.

Our methodology enables construction companies to evaluate the overall project performance with higher accuracy (ie, higher precision) by addressing the indiscrimination problem and higher validity by addressing the redundancy problem (Hand, 2004).

4. Quality assessment of the proposed CI

4.1. Quality assessment overview

We conduct a case study to assess the quality of the proposed methodology for constructing a CI *versus* the alternative normalization and aggregation methods. The test case is the cost performance of 25 real projects provided by a construction company in Korea (Table 1). To evaluate the cost performance of each project, the company measured three sub-indicators: the sales completion rate in percentage, the cost spending rate in percentages, and work productivity in currency (Korean won). The following equations were used in the process:

$$\text{Sales completion rate} = \frac{\text{completed sales}}{\text{planned sales}} = \frac{\text{work quantity completed} \times \text{contracted unit price}}{\text{work quantity planned} \times \text{contracted unit price}} \tag{4}$$

$$\text{Cost spending rate} = \frac{\text{paid cost}}{\text{budget cost}} = \frac{\text{work quantity completed} \times \text{paid unit price}}{\text{work quantity planned} \times \text{budgeted unit price}} \tag{5}$$

Table 1 The three sub-indicators of cost performance in 25 projects

Project	*Sales completion rate (%)*	*Cost spending rate (%)*	*Work productivity (Korean won)*
Pj1	149.10	97.70	25.61
Pj2	100.00	85.60	41.71
Pj3	140.90	100.00	12.59
Pj4	100.00	97.70	16.37
Pj5	100.00	97.40	26.63
Pj6	112.10	99.10	19.04
Pj7	100.00	98.30	27.35
Pj8	134.30	98.50	21.87
Pj9	100.00	99.00	21.00
Pj10	100.00	100.00	12.17
Pj11	100.10	96.80	16.41
Pj12	100.00	97.00	18.00
Pj13	100.00	100.00	23.64
Pj14	100.00	97.30	21.79
Pj15	84.90	99.60	19.50
Pj16	104.20	98.70	18.33
Pj17	118.50	97.40	9.61
Pj18	104.60	97.90	42.95
Pj19	100.00	97.70	25.07
Pj20	102.70	33.70	30.46
Pj21	100.00	99.00	21.79
Pj22	100.00	88.60	38.58
Pj23	105.70	100.00	16.03
Pj24	129.70	97.20	36.62
Pj25	105.40	97.60	23.20

$$\text{Work productivity} = \frac{\text{completed work}}{\text{number of staff}} = \frac{\text{work quantity completed} \times \text{budgeted unit price}}{\text{number of staff}} \quad (6)$$

In this paper, CI quality is assessed using the reliability and accuracy of the measurement results. Measurement reliability is defined as the consistency of a set of measurement results. To assess measurement reliability, we calculate the measurement reliability index of each normalization method (Equation 7) and counted rank inversion.

$$\text{Measurement Reliability Index} = \frac{\Sigma_{i=1}^{n} \text{Max}\{ranks_i\} - \text{Min}\{ranks_i\}}{\textit{The total number of projects}} \quad (7)$$

Here, rank_i represents the rank of measurement results for project i ($i = 1, 2, \ldots, n$), and $\max\{\text{ranks}_i\}$ means the maximum rank among measurement results for project i which are calculated using aggregation methods. Min $\{\text{ranks}_i\}$ means the minimum rank among measurement results for project i which are calculated using aggregation methods.

Intuitively, the larger gap between ranks of a project by each aggregation method means lower reliability. If a normalization method results in a lower reliability index, it might be considered a better reliable normalization method.

Measurement accuracy is defined as the closeness of measured performance results to the actual value of performance. To measure the degree of measurement accuracy, we conduct uncertainty analysis to compare the performance of two projects whose ranks were different based on the five different aggregation methods. We then compare the project performance based on each aggregation method. The uncertainty analysis is implemented in the software *Crystal Ball 11*. We limit ourselves to three types of uncertainties: alternative normalization methods for the values of the sub-indicators; alternative aggregation methods; and uncertainty in the weights of the sub-indicators. Uncertainty analysis focuses on how uncertainty in the input factors propagates through the CI structure and affects its values. Three normalization methods for normalizing each sub-indicator (Z-scores, re-scaling, and the proposed utility function) and five aggregation methods for aggregating normalized sub-indicators (SAW method, WP method, WDI method, TOPSIS and the proposed fuzzy integral) were applied in the present work. Tables 2 and 3 show the normalization and aggregation functions from which the CI could be obtained.

We use the Monte Carlo approach to evaluate the measurement accuracy of the proposed methodology in order to construct CI with K randomly selected input

Table 2 The implementation function for the normalization methods

Method	*Normalization function*		
Standardization	$r_{ij} = \dfrac{x_{ij} - Mean(x_j)}{Stdev(x_j)}$		
Re-scaling	$r_{ij} = \dfrac{x_{ij} - \mathrm{Min}(x_j)}{\mathrm{Max}(x_j) - \mathrm{Min}(x_j)}$		
Utility function	The sales completion rate (r_1):	$r_i = 0$, $r_i = 0.05 \times x_i - 4.25$, $r_i = 1$,	$x_i < 85$ $85 \leqslant x_i \leqslant 105$ $105 < x_i$
	The cost spending rate (r_2):	$r_i = 0$, $r_i = -0.15 \times x_i + 15.24$, $r_i = 1$,	$x_i < 95$ $95 \leqslant x_i \leqslant 101.67$ $101.67 < x_i$
	The work productivity (r_3):	$r_i = 0$, $r_i = 0.1 \times x_i - 1.15$, $r_i = 1$,	$x_i < 11.5$ $11.5 \leqslant x_i \leqslant 21.5$ $21.5 < x_i$

Here, r_{ij} represents the normalized value of the sub-indicator r_j for project i.
i = the project ($i = 1, 2, 3, \ldots, 25$).
j = the sub-indicator ($j = 1, 2, 3$).

Table 3 The implementation function for the aggregation methods

Method	*Aggregation function*
SAW	$CI_i = \Sigma_{j=1}^{n} w_j r_{ij} \quad (i = 1, 2, 3, \ldots, m)$
WP	$CI_i = \Pi_{j=1}^{n} (r_{ij})^{w_j} \quad (i = 1, 2, 3, \ldots, m)$
WDI	$CI_i = \sqrt{\Sigma_{j=1}^{n} (w_j r_{ij})^2} \quad (i = 1, 2, 3, \ldots, m)$
TOPSIS	$CI_i = \dfrac{\sqrt{\Sigma_{j=1}^{n} (w_j r_{ij} - \min_i \{w_j r_{ij}\})^2}}{\sqrt{\Sigma_{j=1}^{n} (w_j r_{ij} - \min_i \{w_j r_{ij}\})^2} - \sqrt{\Sigma_{j=1}^{n} (w_j r_{ij} - \max_i \{w_j r_{ij}\})^2}} \quad (i = 1, 2, 3, \ldots, m)$
Fuzzy integral	$CI_i = \int_x h(x) \circ g(\cdot) = \Sigma_{j=1}^{n} [h(x_{ij}) - h(x_{ij-1})] g(H_i) \quad (i = 1, 2, 3, \ldots, m)$

i = the project ($i = 1, 2, 3, \ldots, 25$).
j = the sub-indicator ($j = 1, 2, 3$).

factors X_1–X_5. The procedure for the Monte Carlo approach follows (Zhou and Ang, 2009):

Step 1: Randomly generate five independent input factors based on the PDF assigned to X_1-X_5, and repeat it K times. That is, generate randomly K combination of five independent input factors $X(t)$, with $t=1,2,\ldots,K$ (a $X(t)$ sets of input factors generated as $X_1(t)$, $X_2(t), \ldots, X_5$ (t) $(t=1,2,\ldots,K)$).

Step 2: For each set of input factors $X_1(t)-X_5(t)$ $(t=1,2,\ldots,K)$, use the disposal rule defined in Table 4 to select the corresponding normalization and aggregation methods, and determine the weights for the three sub-indicators.

Step 3: For $t=1,2,\ldots,K$, use the normalization and aggregation methods assigned to derive the corresponding CI. Data for each sub-indicator are first normalized according to the trigger X_1 that is sampled from a uniform distribution [0, 1), where $0 \leqslant X_1 < (1/3)$ is used for re-scaling, $(1/3) \leqslant X_1 < (2/3)$ is used for standardization and $(2/3) \leqslant X_1 < 1$ is used for the utility function. Second, the data for each normalized sub-indicator are aggregated according to the trigger X_2. The trigger X_2, with the same type of PDF as X_1, guides the selection of an aggregation method. Finally, in the case of SAW, WP, WDI, TOPSIS, the three values from independent uniform $[0,1]$ distributions are scaled to a unit sum in order to obtain w_1-w_3. On the other hand, in the case of fuzzy integral, the degree of influence of each sub-indicator can be determined without consideration of the constraint that the sum of these values must be one (eg $w_1=0.3$, $w_2=0.6$, and $w_3=0.5$). Therefore, the three values are selected from independent uniform $[0,1]$ distributions such as the disposal rule in Table 4.

Step 4: Analyse the results to assess measurement accuracy.

4.2. *Quality assessment results*

4.2.1. Measurement reliability. To assess the measurement reliability, we calculate the reliability index using Equation (7) and count rank inversion based on each normalization method. Our results show that the reliability of the proposed normalization method (utility function) is higher than that of the two different normalization methods (Table 5).

4.2.2. Measurement accuracy. To assess measurement accuracy, we conduct an uncertainty analysis to compare the performance of two projects whose ranks differ based on the five different aggregation methods. We then compare the project performance based on each aggregation method.

Table 4 The five uncertain input factors

Input factor	*Definition*	*PDF*	*Disposal rule*
X_1	Trigger to select normalization method	Uniform [0, 1)	$[0,1/3) \equiv Z-\text{score}, [1/3,2/3) \equiv \text{Re}-\text{scalling}, [1/3,2/3) \equiv \text{Utility function}$
X_2	Trigger to select aggregation method	Uniform [0, 1)	$[0,02) \equiv \text{SAW}, [0.2,0.4) \equiv \text{WP}, [0.4,0.6) \equiv \text{WDI}, [0.6,0.8) \equiv \text{TOPSIS}\ [0.8,1) \equiv \text{Fuzzy Integral}$
X_3	w_1	Uniform [0, 1]	$w_1 = \frac{X_3}{\Sigma_{k=3}^5 X_k}$ (SAW, WP, WDI, TOPSIS) $w_1 = X_3$ (Fuzzy Integral)
X_4	w_2	Uniform [0, 1]	$w_2 = \frac{X_4}{\Sigma_{k=3}^5 X_k}$ (SAW, WP, WDI, TOPSIS) $w_2 = X_4$ (Fuzzy Integral)
X_5	w_3	Uniform [0, 1]	$w_3 = \frac{X_5}{\Sigma_{k=3}^5 X_k}$ (SAW, WP, WDI, TOPSIS) $w_3 = X_5$ (Fuzzy Integral)

Table 5 Comparison of the normalization method results

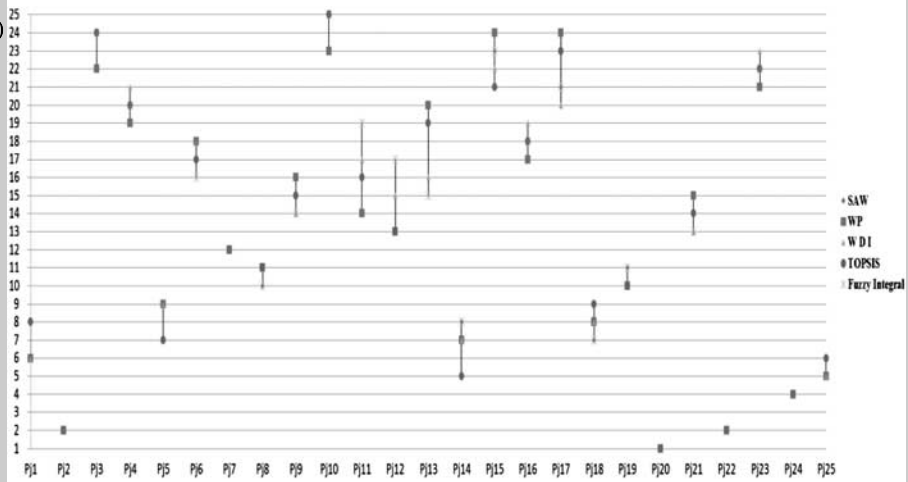
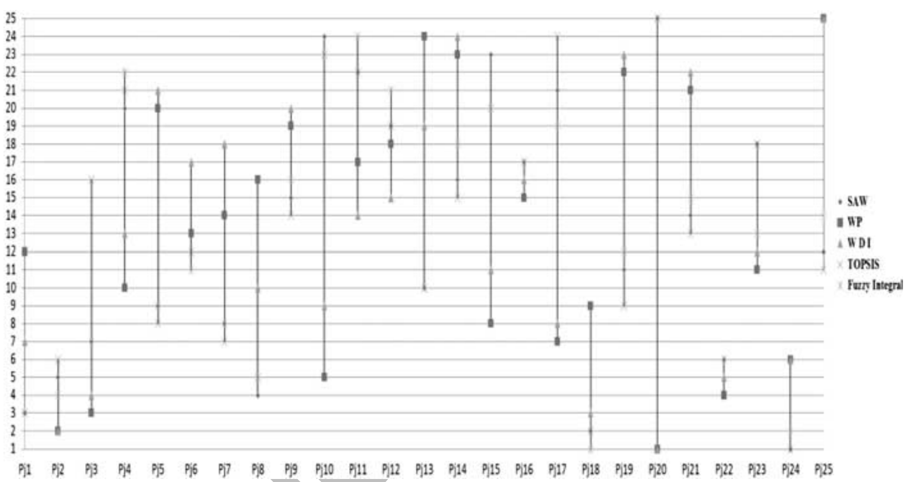
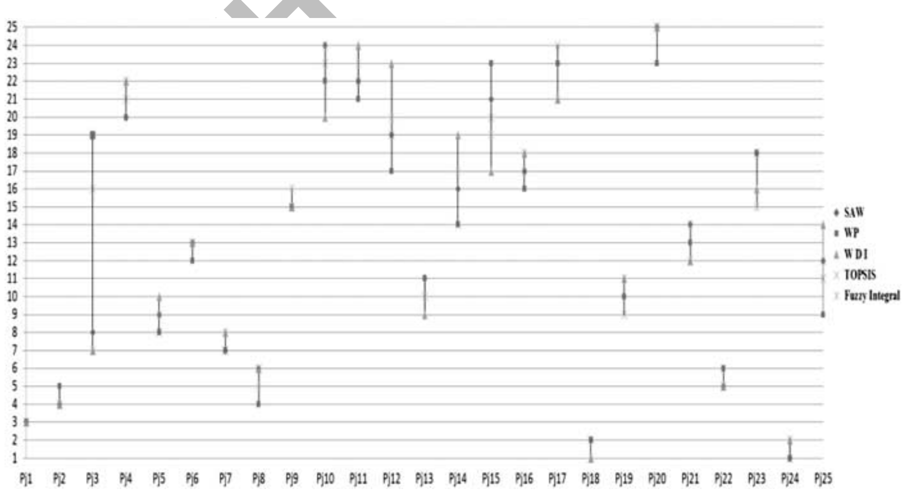
<table>
<tr><th></th><th>Comparing normalization method</th><th>Measurement reliability</th><th>Rank inversion</th></tr>
<tr><td>Utility function (Proposed method)</td><td>
</td><td>1.96</td><td>26</td></tr>
<tr><td>Z-Score</td><td>
</td><td>10.44</td><td>178</td></tr>
<tr><td>Re-scaling</td><td>
</td><td>2.80</td><td>41</td></tr>
</table>

Table 6 Comparison of the project performance rank (with utility function as the normalization method)

Fuzzy integral (Proposed method)	*SAW*	*WP*	*WDI*	*TOPSIS*
Pj1 > Pj14	Pj1 > Pj14	Pj1 > Pj14	Pj1 > Pj14	Pj1 < Pj14
Pj6 > Pj16	Pj6 > Pj16	Pj6 < Pj16	Pj6 > Pj16	Pj6 > Pj16
Pj11 < Pj16	Pj11 > Pj16	Pj11 > Pj16	Pj11 > Pj16	Pj11 > Pj16
Pj12 < Pj21	Pj12 > Pj21	Pj12 > Pj21	Pj12 < Pj21	Pj12 > Pj21
Pj14 < Pj18	Pj14 < Pj18	Pj14 > Pj18	Pj14 > Pj18	Pj14 > Pj18
Pj15 < Pj23	Pj15 < Pj23	Pj15 < Pj23	Pj15 > Pj23	Pj15 > Pj23

Underline means the comparison results by proposed method equal comparison results by alternative method.

We conduct uncertainty analysis targeting six of 26 rank inversion examples (Table 6). The histogram shown in Table 7 represents the outcome of the uncertainty analysis on the differences in the CI values between the two projects. (a) Comparing performance between project 1 and project 14 (first figure in Table 7): the right-hand region, in which project 1 performs better than project 14, covers about 93.862% of the total area. (b) Comparing performance between project 6 and project 16 (second figure in Table 7): the right-hand region, in which project 6 performs better than project 16, covers about 85.816% of the total area. (c) Comparing performance between project 11 and project 16 (third figure in Table 7): the left-hand region, in which project 16 performs better than project 11, covers about 77.958% of the total area. (d) Comparing performance between project 12 and project 21 (fourth figure in Table 7): the left-hand region, in which project 21 performs better than project 12, covers about 75.303% of the total area. (e) Comparing performance between project 14 and project 18 (fifth figure in Table 7): the left-hand region, in which project 18 performs better than project 14, covers about 89.625% of the total area. (f) Comparing performance between project 15 and project 23 (sixth figure in Table 7): the left-hand region, in which project 23 performs better than project 15, covers about 72.959% of the total area.

Next, we conduct uncertainty analysis to compare project performance based on each aggregation method whose rank is equal to the result using the fuzzy integral. (a) Comparing performance between project 1 and project 14, project 1 performs better and covers a larger area with the fuzzy integral than that covered by SAW, WP, and WDI (96.203, 95.958, 94.421, and 93.155%, respectively. (b) Comparing projects 6 and 16, project 6 performs better and covers a larger area with the fuzzy integral than that covered by SAW, WDI, and TOPSIS (fuzzy integral > SAW > TOPSIS > WDI; 95.703, 93.280, 86.640, and 75.505%, respectively). (c) Comparing projects 11 and 16, project 16 performs better, covering about 77.958% of the total area. (d) Comparing projects 12 and 21, project 21 performs better and covers a larger area with the fuzzy integral than that covered by WDI (fuzzy integral > WDI; 91.609 and 61.726%, respectively). (e) Comparing projects 14 and 18, project 18 performs better and covers a larger area with the fuzzy integral than that covered by SAW (fuzzy integral > SAW; 94.034 and 93.442%, respectively). (f) Comparing project 15 and project 23, project 23 performs better and covers a larger area with the fuzzy integral than that covered by SAW, and WP (fuzzy integral > SAW > WP; 84.930, 83.972, and 71.441%respectively). As a result, the measurement accuracy of the proposed methodology is greater than that of the four different aggregation methods (Table 8).

Our results show that the proposed methodology helps evaluate the overall project performance with a higher degree of measurement reliability and accuracy compared with the alternative methodology.

5. Conclusion and further research

The conventional methodology of the overall project performance evaluation in construction organizations uses a categorical scale, budget allocation, and simple additive aggregation function. Combined with the characteristics of sub-indicators of construction projects, this set of methods causes indiscrimination and redundancy. Although many methods for normalization, weighting, and aggregation exist for developing a CI, an appropriate set of methods that address these problems are yet to be developed. To address these problems in evaluating the overall project performance evaluation, we propose a novel methodology that utilizes fuzzy theories. It includes the following three elements: (1) a utility function for normalizing the values of sub-indicators, (2) a fuzzy measure for weighting the sub-indicators, and (3) a fuzzy integral for aggregating the values of the sub-indicators.

To demonstrate its suitability, we assessed the quality of the proposed methodology for constructing CIs in comparison to conventional methodologies. In this paper, CI quality was assessed using measurement reliability and measurement accuracy. We calculated the measurement reliability index using each normalization method. The result shows that the measurement reliability of the proposed method (1.96) is greater than that of the two

Table 7 Project performance comparison results

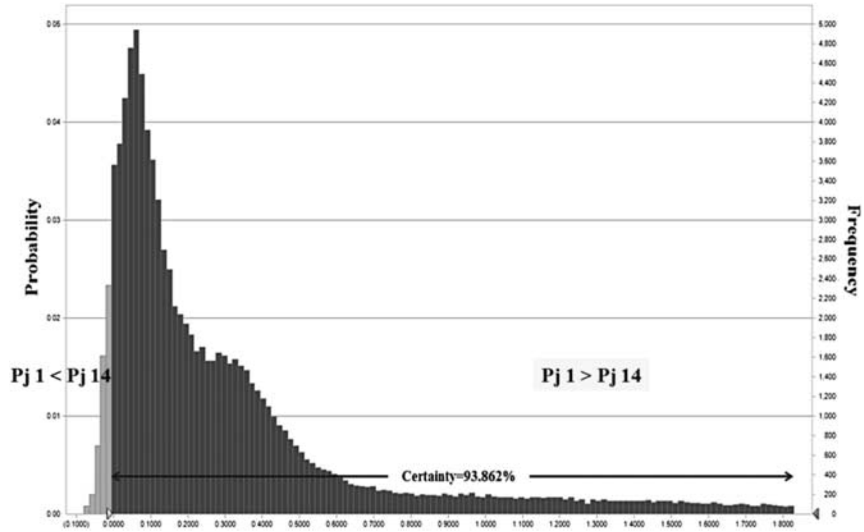
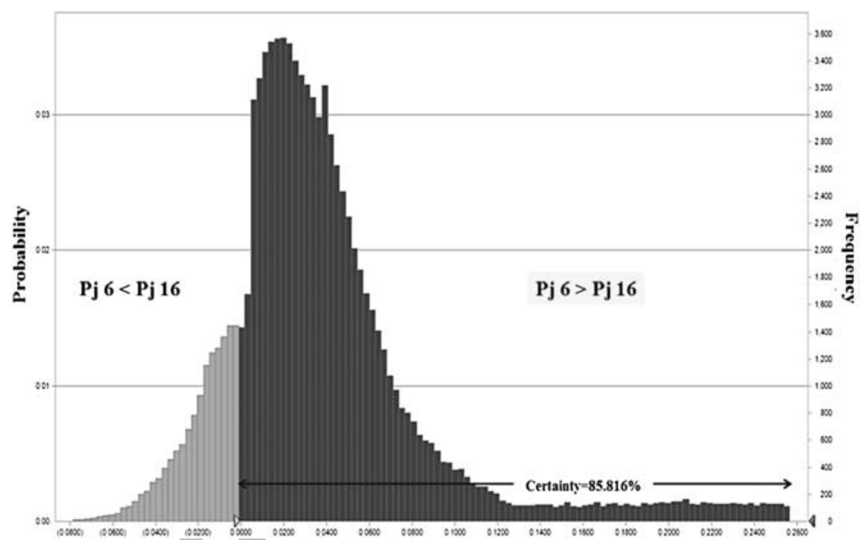
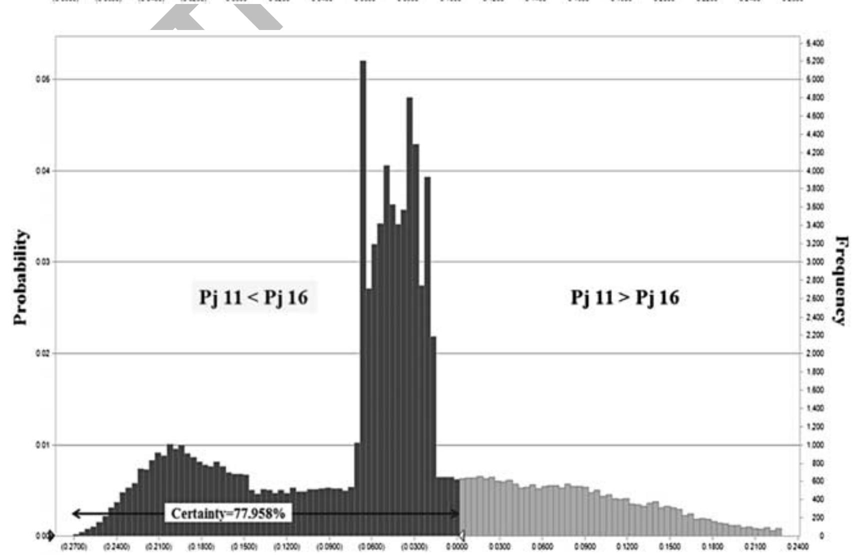
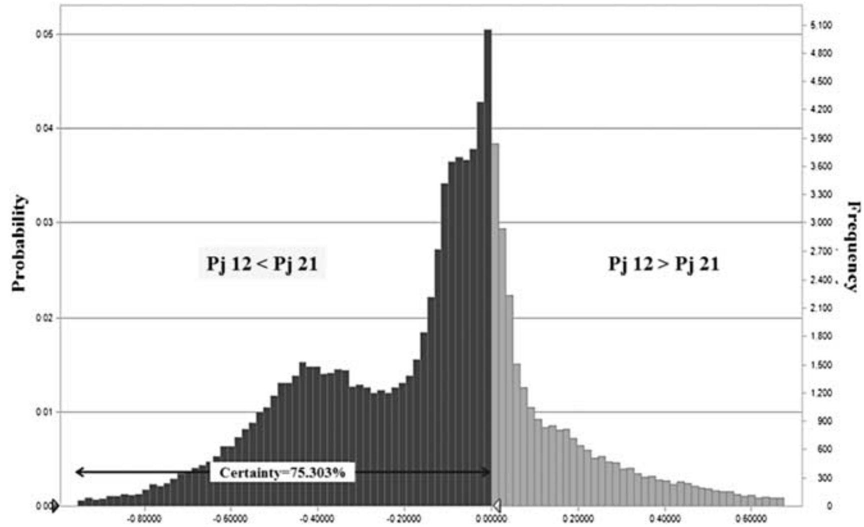
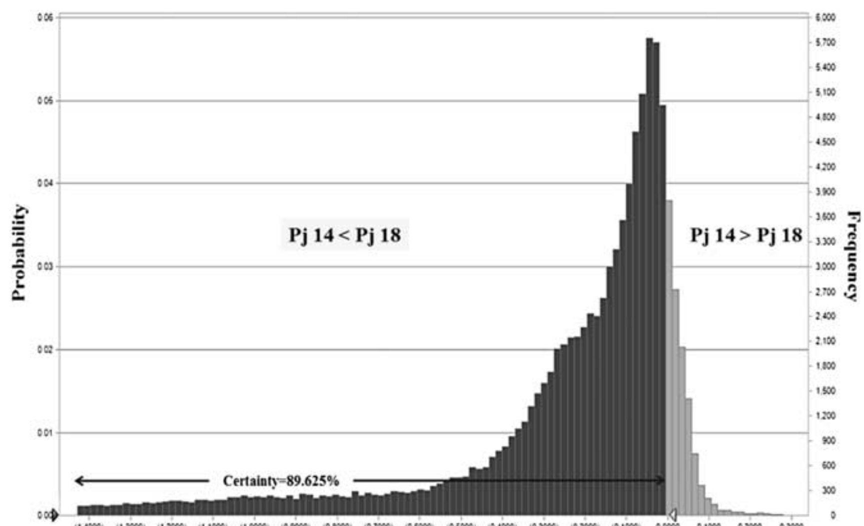
	Comparing project performance	*Certainty (%)*
Pj1 > Pj14		93.862
Pj6 > Pj16		85.816
Pj11 < Pj16		77.958

Table 7 *Continued*

Pj12 < Pj21 75.303



Pj14 < Pj18 89.625



Pj15 < Pj23 72.959

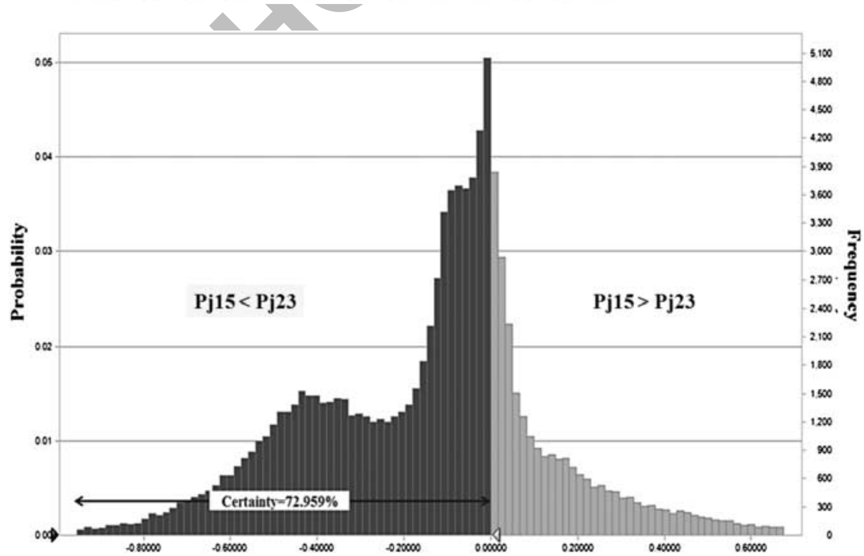


Table 8 Aggregation method comparison results

Comparing project performance	*Aggregation method*	*Certainty (%)*
Pj1 > Pj14	Fuzzy integral (Proposed method)	96.203
	SAW	95.985
	WP	93.155
	WDI	94.421
Pj6 > Pj16	Fuzzy integral (Proposed method)	95.730
	SAW	93.280
	WDI	75.505
	TOPSIS	86.640
Pj11 < Pj16	Fuzzy integral (Proposed method)	77.958
Pj12 < Pj21	Fuzzy integral (Proposed method)	91.609
	WDI	61.726
Pj14 < Pj18	Fuzzy integral (Proposed method)	94.034
	SAW	93.442
Pj15 < Pj23	Fuzzy integral (Proposed method)	84.930
	SAW	83.972
	WP	71.441

different normalization methods (10.44 and 2.8, respectively). We conducted uncertainty analysis to compare the performance of two projects whose ranks based on the five different aggregation methods were different. We then compared project performance based on each aggregation method. Our results show that the measurement accuracy of the proposed methodology is greater than that of the four different aggregation methods. Therefore, the proposed methodology significantly improves the reliability and accuracy of project performance. That is, with our proposed methodology, construction companies can more consistently and accurately evaluate the overall project performance or project success.

Although this research used real project performance data, only three sub-indicators related to cost performance on 25 projects were tested. Future research to expand the number of projects and include qualitative sub-indicators is required, along with taking into account different project characteristics and investigating under- and over-estimated projects in depth. In addition, sensitivity analysis is required to analyse the degree to which each individual source of uncertainty contributes to output variance.

Acknowledgements—This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (2012-0005376). The present research has been conducted by the Research Grant of Kwangwoon University in 2012.

References

Bai J, Yang X and Tao L (2011). Research on construction project process performance measurement. In: *2011 IEEE 18th International Conference on Industrial Engineering and Engineering Management*, Institute of Electrical and Electronics Engineers (IEEE): Changchun, China, pp 1915–1918.

Cha H and Kim C (2011). Quantitative approach for project performance measurement on building construction in South Korea. *KSCE Journal of Civil Engineering* **15**(8): 1319–1328.

Chiu YC, Shyu JZ and Tzeng GH (2004). Fuzzy MCDM for evaluating the E-commerce strategy. *International Journal of Computer Applications in Technology* **19**(1): 12–22.

Clivillé V, Berrah L and Mauris G (2007). Quantitative expression and aggregation of performance measurements based on the MACBETH multi-criteria method. *International Journal of Production Economics* **105**(1): 171–189.

Dainty A, Cheng MI and Moore D (2003). Redefining performance measures for construction project managers: An empirical evaluation. *Construction Management and Economics* **21**(2): 209–218.

Diaz-Balteiro L and Romero C (2004). In search of a natural systems sustainability index. *Ecological Economics* **49**(3): 401–405.

Ebert U and Welsch H (2004). Meaningful environmental indices: A social choice approach. *Journal of Environmental Economics and Management* **47**: 270–283.

Esty DC, Levy MA, Srebotnjak T and de Sherbinin A (2005). *2005 Environmental Sustainability Index: Benchmarking National Environmental Stewardship*. Yale Center for Environmental Law & Policy: New Haven, CT, USA.

Fayek AR and Sun Z (2001). A fuzzy expert system for design performance prediction and evaluation. *Canadian Journal of Civil Engineering* **28**(1): 1–25.

Fayek AR and Oduba A (2005). Predicting industrial construction labor productivity using fuzzy expert systems. *Journal of Construction Engineering and Management* **131**(8): 938–941.

Freudenberg M (2003). *Composite indicators of country performance: A critical assessment*. OECD Science, Technology and Industry Working Papers, OECD, Directorate for Science, Technology and Industry, http://dx.doi.org/10.1787/405566708255.

Grabisch M (1996). The application of fuzzy integrals in multicriteria decision making. *European Journal of Operational Research* **89**(3): 445–456.

Hand DJ (2004). *Measurement Theory and Practice: The World Through Quantification*. John Wiley & Sons Ltd: New York.

Jacobs RP and Goddard M (2004). *Measuring performance: An examination of composite performance indicators*. Centre for Health Economics, Technical Paper Series 29.

Kumaraswamy MM and Thorpe A (1996). Systematizing construction project evaluations. *Journal of Management in Engineering* **12**(1): 34–39.

Landy F and Farr J (1983). *The Measurement of Work Performance: Methods, Theory, and Applications*. Academic Press: New York.

Lauras M, Marques G and Gourc D (2010). Towards a multidimensional project performance measurement system. *Decision Support Systems* **48**(2): 342–353.

Lun G, Holzer D, Tappeiner G and Tappeiner U (2006). The stability of rankings derived from composite indicators: Analysis of the 'Il Sole 24 Ore' quality of life report. *Social Indicators Research* **77**(2): 307–331.

Liginlal D and Ow T (2006). Modeling attitude to risk in human decision processes: An application of fuzzy measures. *Fuzzy Sets and Systems* **157**(23): 3040–3054.

Marques G, Gourc D and Lauras M (2010). Multi-criteria performance analysis for decision making in project management. *International Journal of Project Management* **29**(8): 1057–1069.

Nardo M, Saisana M, Saltelli A and Tarantola S (2005). *Tools for Composite Indicators Building*. European Commission-Joint Research Centre: Ispra, VA, Italy.

OECD and European Commission-Joint Research Centre (2008). *Handbook on Constructing Composite Indicators: Methodology and user Guide*. OECD Publishing: Paris, France.

Park M, Kim N, Lee H, Ahn C and Lee K (2009). Construction project performance management using BSC and data warehouse. *Journal of Korean Institute of Construction Engineering and Management* **10**(2): 14–25.

Saltelli A (2006). Composite indicators between analysis and advocacy. *Social Indicators Research* **81**(1): 65–77.

Saisana M and Tarantola S (2002). *State-of-the-art Report on Current Methodologies and Practices for Composite Indicator Development*. European Commission-Joint Research Centre: Ispra, VA, Italy.

Sinha BK and Shah KR (2003). On some aspects of data integration techniques with environmental applications. *Environmetrics* **14**(4): 409–416.

Tzeng G, Ouyang Y, Lin C and Chen C (2005). Hierarchical MADM with fuzzy integral for evaluating enterprise intranet web sites. *Information Sciences* **169**(3–4): 409–426.

Yoon KP and Hwang CL (1995). *Multiple Attribute Decision Making: An Introduction*. Sage Publications: Thousand Oaks, CA.

Zhou P, Ang BW and Poh KL (2006). Comparing aggregating methods for constructing the composite environmental index: An objective measure. *Ecological Economics* **59**(3): 305–311.

Zhou P and Ang W (2009). Comparing MCDA aggregation methods in constructing composite indicators using the Shannon-Spearman measure. *Social Indicators Research* **94**(1): 83–96.

Received March 2012;
accepted January 2013 after three revisions