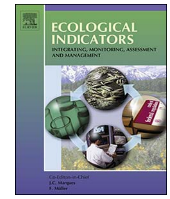
Ecological Indicators 81 (2017) 491–502

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind



Review

# When to use what: Methods for weighting and aggregating sustainability indicators

Xiaoyu Gan[a], Ignacio C. Fernandez[b], Jie Guo[c,*], Maxwell Wilson[d], Yuanyuan Zhao[e], Bingbing Zhou[b], Jianguo Wu[b,d,f]

[a] College of Architecture and Environment, Sichuan University, Chengdu, Sichuan Province, China
[b] School of Sustainability, Arizona State University, Tempe, AZ, USA
[c] College of Land Management, Nanjing Agricultural University, Nanjing, Jiangsu Province, China
[d] School of Life Sciences, Arizona State University, Tempe, AZ, USA
[e] School of Soil and Water Conservation, Beijing Forestry University, Beijing, China
[f] Center for Human-Environment System Sustainability (CHESS), State Key Laboratory of Earth Surface Processes and Resource Ecology (ESPRE), Beijing Normal University, Beijing 100875, China

A R T I C L E I N F O

*Keywords:*
Aggregation
Indicator framework
Sustainability indicators
Sustainability indices
Weighting

A B S T R A C T

*Context:* Sustainability indices (SIs) have become increasingly important to sustainability research and practice. However, while the validity of SIs is heavily dependent on how their components are weighted and aggregated, the typology and applicability of the existing weighting and aggregation methods remain poorly understood.
*Objectives:* To close the knowledge gap regarding when to use which weighting and aggregation methods for constructing SIs, we review the most commonly used methods for weighting and aggregating SIs, discuss their benefits and drawbacks, and suggest a process-oriented approach for choosing appropriate weighting and aggregation methods depending on research objectives.
*Methods:* Our review synthesis was based on peer-reviewed journal articles, books, and reports by international organizations, governmental agencies, and research institutions. After carefully examining their principles, characteristics, and applications, we selected and classified the frequently used methods for indicator weighting and aggregation.
*Results:* We systematically discuss the benefits and drawbacks of nine weighting methods and three aggregation methods. We propose a four-step process for choosing the most suitable weighting and aggregation methods based on: research purposes, spatial and temporal scales, and sustainability perspectives.
*Conclusions:* In this research, we chose the most commonly used methods for weighting and aggregating SIs and analyzed the characteristics, strengths and weaknesses of each method. We found that choosing appropriate weighting and aggregation methods for a specific sustainability assessment project is an extremely important and challenging task. To meet this challenge, we propose a process-oriented approach for properly selecting methods according to the purpose, scale and sustainability concept. This approach can facilitate the proper selection of these methods in sustainability research and practice.

## 1. Introduction

Sustainability is the challenge of our time (Sachs, 2015). By seeking to achieve dynamic and simultaneous harmony among ecological subsystems (environmental sustainability), social subsystems (social sustainability), and economic subsystems (economic sustainability), sustainability is inherently complex, multi-dimensional, and embedded with trade-offs among multiple sustainability dimensions (Wu, 2013). However, as the public's desire for more sustainability grows stronger (Kates and Clark, 1999; Kates et al., 2001), so does the need to accurately assess the sustainability of our societies (Böhringer and Jochem, 2007), which is no easy task. To capture the complexity of sustainability, sustainability assessments often require the integration of multiple indicators to form composite indices (Özdemir et al., 2011; Wu and Wu, 2012). Thus, while developing sustainability indicators and indices (SIs) is a critical tool for assessing and ultimately attaining sustainability, the specifics of SI formulation can radically impact the measured sustainability of a system (Singh et al., 2009; Wilson and Wu, 2017).

The main procedures for building a sustainability index include

* Corresponding author at: College of Land Management, Nanjing Agricultural University, #1 Weigang Road, Nanjing 210095, Jiangsu Province, China.
*E-mail address:* guojie@njau.edu.cn (J. Guo).

http://dx.doi.org/10.1016/j.ecolind.2017.05.068
Received 8 February 2017; Received in revised form 9 May 2017; Accepted 29 May 2017
Available online 29 June 2017

selecting appropriate sustainability indicators, weighting the selected indicators, and aggregating those indicators into a composite index (Meadows, 1998; Juwana et al., 2012). Disagreements on indicator selection are relatively easy to decipher, as existing guidelines, e.g., Bellagio Principles (Hardi and Zdan, 1997), or indicator frameworks, such as the Pressure-State-Response framework (OECD, 1993), can provide guidance for indicator selection (Wu and Wu, 2012). However, because the process of indicator integration is an inherently subjective procedure (Morse et al., 2001), selecting appropriate weighting and aggregation methods is challenging (Saisana and Tarantola, 2002; Wilson and Wu, 2017).

The weighting and aggregation of index components are critically important steps in any sustainability assessment. Weights of SIs reflect the relative importance of different dimensions in their contributions to the sustainability performance of a system, while aggregation essentially reflects the substitutability of different dimensions. The weighting and aggregation methods utilized in SI formulation define whether dimensions can compensate or substitute for each other. Whether complete, partial, or no substitution between environmental (or natural) and socioeconomic capital is legitimate underscores the two widely discussed sustainability perspectives: weak sustainability and strong sustainability (Daly et al., 1995; Markulev and Long, 2013). Weak sustainability allows for unlimited substitution between sustainability dimensions. Strong sustainability is a paradigm that views economic activities as part of the social domain, and both economic and social actions are constrained by the environment (Wu, 2013). Each perspective dictates a different set of criteria for indicator selection and fundamentally influences the final verdict of a sustainability assessment (Wu, 2013; Huang et al., 2015). Further, the weights of SIs not only reflect the relative importance of different dimensions in their contributions to overall sustainability but also symbolize the trade-off ratios among the dimensions if they are conceived as substitutable. Thus, it should come as no surprise that the inappropriate selection of weighting or aggregating methods can cause SIs to provide misleading information (Böhringer and Jochem, 2007). In this sense, one of the main challenges in developing and applying SIs is to know “when to use what.”

Informative reviews have been published on the strengths and weaknesses of commonly used sustainability indices (Böhringer and Jochem, 2007; Mayer, 2008; Singh et al., 2009; Mori and Christodoulou, 2012). These studies provided suggestions on how to choose appropriate sustainability indicators and indices for policy decisions and discussions of those SIs’ formulation and performance. Researchers have also proposed guidelines for constructing sustainability indices in various contexts, such as urban sustainability (Huang et al., 2015), industry sustainability (Singh et al., 2007), energy sustainability (Wang et al., 2009), and agricultural sustainability (Gómez-Limón and Sanchez-Fernandez, 2010). The main goal of the present study, therefore, is to provide a practical guide for the selection of weighting and aggregation methods in the formulation and application of SIs. We focus on three main questions: (1) What are the most commonly used methods for weighting and aggregation reported in the literature? (2) What are the strengths and weaknesses of these methods for measuring sustainability? (3) How can these methods be best utilized in SI development and applications? To address these questions, we systematically reviewed the main methods for weighting and aggregating SIs, identified the main advantages and challenges for applying these methods, and proposed a process-oriented approach to help researchers and practitioners select the most suitable weighting and aggregation methods for sustainability assessment using SIs in different contexts.

## 2. Methodology

### 2.1. Analytical framework

To ensure this review is as representative as possible, an analytical framework in which weighting and aggregation methods used for constructing SIs was needed. In our paper, the classification strategy of weighting and aggregation methods proposed by Nardo et al. (2005) and OECD (2008) was adopted.

Within this framework, methods for weighting indicators can be broadly categorized into three main groups: (1) equal weighting, (2) statistic-based weighting, and (3) public/expert opinion-based weighting. Equal weighting means that all the indicators are given the same weight. Statistic-based weighting derives weights from the statistical characteristics of the data (OECD, 2008). Unlike equal weighting and statistic-based weighting, public/expert opinion-based weighting relies on inputs from the participating public or experts, whose judgments ultimately determine the weights to be assigned to individual indicators (OECD, 2008). Thus, weights determined by public/expert opinion reflect the value judgments of the participants regarding different aspects of sustainability (e.g., relative importance, relative urgency, or substitution rates).

In contrast, aggregation methods integrate weighted components (e.g., indicators) into a single composite index. Different classification schemes for aggregation methods exist. In general, classification schemes include those based on the semantics of aggregation (Beliakov et al., 2007; Grabisch, 2009) and those based on the degree of permission of compensation (OECD, 2008). We adopt the latter classification scheme because it has a closer relation to the technical challenges of integrating weighted indicators based on sustainability concepts (Wilson and Wu, 2017). Widely used aggregation methods based on this classification scheme include additive aggregation methods (e.g., arithmetic), multiplicative aggregation methods (e.g., geometric) and non-compensatory aggregation methods (e.g., multi-criteria analysis).

### 2.2. Literature analysis

To evaluate what are the most commonly used methods for weighting and aggregation reported in the literature, we did a statistical analysis of published literature separate from the review discussed in the main text. To select the papers for this analysis, we followed the PRISMA flowchart (Liberati et al., 2009), shown in Fig. 1. We searched papers using the Web of Knowledge database by using the search by topic option with the search terms shown in Table 1. This search was done on April, 14$^{th}$, 2017, resulting in 1319 publications, after removing duplicates. We added to these publications 98 documents that were considered relevant, but were not available on the Web of Knowledge database. We selected these publications based on the references from Nardo et al. (2005), Böhringer and Jochem (2007), Singh et al. (2009), and Huang et al. (2015). Titles and abstracts of these papers (n = 1417) were then screened to remove: (1) papers that were cited less than 30 times, (2)literature that was unrelated to sustainability assessments, and (3) papers on indicator sets instead of composite indicators. The remaining 230 articles were then assessed to remove articles presenting indices that did not provide specific weighting and aggregating methods or that provided duplicate indices without any modifications in the methods used for weighting or aggregation. A total of 90 SIs, including 96 weighting scheme variations and 90 aggregation scheme variations, were identified. As some SIs use different weighting/aggregation methods to integrate sub-indicators into the final composite indices, we counted each weighting/aggregation scheme as a separate index, and thus a total of 96 different SIs were used for the analysis.

### 2.3. Literature analysis results

Among the 96 SIs reviewed in our paper, 46.88% adopted equal weighting methods, 21.88% adopted statistical-based methods (principal component analysis, benefit of the doubt approach, regression analysis, unobserved component models), and 23.95% adopted