European Journal of Operational Research 264 (2018) 1–16

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Invited Review

Data envelopment analysis application in sustainability: The origins, development and future directions

Haibo Zhou[a], Yi Yang[b], Yao Chen[b], Joe Zhu[c,*]

[a] School of Economics and Management, Southeast University, Nanjing, Jiangsu 210096, PR China
[b] Manning School of Business, University of Massachusetts at Lowell, Lowell, MA 01845, USA
[c] Foisie Business School, Worcester Polytechnic Institute, Worcester, MA 01609, USA

ARTICLE INFO

Article history:
Received 12 September 2016
Accepted 7 June 2017
Available online 15 June 2017

Keywords:
Data envelopment analysis (DEA)
Sustainability
Literature survey
Citation analysis

ABSTRACT

Sustainable development and sustainability assessment have been of great interest to both academe and practitioners in the past decades. In this study, we review the literature on data envelopment analysis (DEA) applications in sustainability using citation-based approaches. A directional network is constructed based on citation relationships among DEA papers published in journals indexed by the Web of Science database from 1996 to March 2016. We first draw the citation chronological graph to present a complete picture of literature development trajectory since 1996. Then we identify the local main DEA development paths in sustainability research by assigning an importance index, namely search path count (SPC), to each link in the citation network. The local main path suggests that the current key route of DEA applications in sustainability focus on the environmental sustainability. Through the Kamada–Kawai layout algorithm, we find four research clusters in the literature including corporate sustainability assessment, regional sustainability assessment, sustainability composite indicator construction, and sustainability performance analysis. For each of the clusters, we further identify the key articles based on citation network and local citation scores, demonstrate the developmental trajectory of the literature, and suggest future research directions.

1. Introduction

The concept of sustainability stems from ecology and describes the use of a regenerative natural system in such a way that this system retains its essential properties and its population can naturally be replenished. In more general terms, sustainability is the endurance of systems and processes. The organizing principle for sustainability is sustainable development, which finds its way into the economics and management areas in 1987 when sustainable development was first initiated as an environmentally friendly, economically feasible and socially acceptable growth pattern in the Brundtland Commission (formally named as the World Commission on Environment and Development (WCED). Since then, thousands of initiatives have been taken at the local, national, and global levels in an attempt to address different aspects of the sustainability challenges (Mebratu, 1998).

Since the early 2000s, firms have been pressured to pay attention to the triple bottom line of sustainability – profit, people and planet (Elkington, 2002) because of the increasing demand for natural resources (clean water, crude oil, woods, metals, etc.) whose supply continues to diminish, the raised concerns about various unethical corporate practices and the development of the emerging markets with supply-chain constraints (Tang & Zhou, 2012). As a result, the need for measuring sustainable development is widely recognized (e.g., Tyteca, 1998). So far, sustainability assessment has served four major purposes: decision making and management, advocacy, participation and consensus building, and research and analysis (Parris & Kates, 2003), and been applied at different levels: national (e.g., Coli, Nissi, & Rapposelli, 2011; Munksgaard, Wier, Lenzen, & Dey, 2005), regional or urban community (e.g., Hu, Sheu, & Lo, 2005; Munda & Saisana, 2011), industry sectorial (e.g., Peres-Neto, Legendre, Dray, & Borcard, 2006; Zofio & Prieto, 2001), and corporate (e.g., Figge & Hahn, 2004; Kuosmanen & Kuosmanen, 2009). In the beginning, sustainability assessment mainly focused on environmental sustainability problems covering only economic and environmental dimensions. More recently, this line of research has started to focus on prospects for lasting net gains and the acceptability of trade-off rules among the environmental, economic and social dimensions (Gibson, 2006; Pope, Annandale, & Morrison-Saunders, 2004; Winfield, Gibson, Markvart, Gaudreau, & Taylor, 2010).

* Corresponding author.
E-mail addresses: zhouhb0920@126.com (H. Zhou), Yi_Yang@uml.edu (Y. Yang), Yao_Chen@uml.edu (Y. Chen), jzhu@wpi.edu (J. Zhu).

http://dx.doi.org/10.1016/j.ejor.2017.06.023
0377-2217/

Accordingly, three categories of indicators and methods for sustainability evaluation have emerged in the literature. *System analysis* is an approach that takes in consideration of both the relationships between the internal components of the system, and the relationships between internal components and external environment (e.g., Antonio, Cristina, & Stefano, 2012; Goerner, Lietaer, & Ulanowicz, 2009; Ulanowicz, 2009). *Flow analysis* evaluates system sustainability through resource utilization efficiency that only considers the relationship between internal components and the external environment (e.g., Balocco, Papeschi, Grazzini, & Basosi, 2004; Campbell & Garmestani, 2012; Paoli, Vassallo, & Fabiano, 2008). Finally, *indicator enumeration* mainly chooses indicators from environmental, economic, social and institutional aspects to evaluate the system sustainability without considering either of the relationships mentioned above (e.g., Ness, Urbel-Piirsalu, Anderberg, & Olsson, 2007; Ou & Liu, 2010; Yli-Viikari, 1999).

Data Envelopment Analysis (DEA) (Charnes, Cooper, & Rhodes, 1978) is a method for evaluating performance of peer decision making units (DMUs) with multiple performance measures that are termed as inputs and outputs. DEA first establishes an 'efficient frontier' formed by a set of DMUs that exhibit best practices and then assigns the efficiency level to other non-frontier units according to their distances to the efficient frontier. Over the years, DEA has been enriched and modified. Numerous DEA models have been developed and used in various applications including sustainability research. In general, there are three approaches to employ DEA models in the sustainability literature (Choi & Zhang, 2011): traditional DEA models with simple translation of data (Lovell, Pastor, & Turner, 1995; Yeh, Chi, & Hsu, 2010), traditional DEA models treating undesirable outcomes as inputs (Hu & Wang, 2006; Zhang, Bi, Fan, Yuan, & Ge, 2008), and DEA models employing the concept of weak disposability technology (Färe & Grosskopf, 2004; Zhou, Ang, & Poh, 2008a). Researchers have applied DEA models to address corporate, regional and national sustainability issues as well as those related to supply chain.

Although DEA has been extensively applied in sustainability, few surveys to the best of our knowledge have been conducted to systematically review the current status of the literature and discuss the future research direction except for Dakpo, Jeanneaux, and Latruffe (2016). Although Dakpo et al. (2016) make a critical review on methods integrating environmental aspects into productive efficiency, their study focuses on only environmental factors, especially the undesirable outputs in production technology modeling, and does not include social factors, another important part of sustainability. In addition, their review is based on subjective and qualitative analyses rather than objective quantitative analysis methods. To fill the gap, our study collects 320 relevant papers published from 1996 to March 2016 and analyzes the research status of DEA applications in sustainability through citation analysis of bibliometrics. Using the citation analysis software HistCite, we conduct a visual analysis and construct a citation chronological graph to identify the main development route and key publications of DEA application in sustainability. Then with the help of Pajek software, we discover the major research clusters as well as the local main paths, and further identify future research directions in each research cluster. In addition, our review highlights the importance of reliable sustainability measures and introduces current major DEA approaches in sustainability evaluation.

This paper is organized as follows. In the next section, we describe the data and methods used in this study. Section 3 discusses the basic statistics for the DEA applications in sustainability. Section 4 presents the major findings through citation chronological graph and main path analysis. Section 5 identifies the major research clusters and draws the development trajectories of each cluster, which presents the most-cited works in each research area. The last section draws conclusions including implications and insights from the analysis results.

2. Review methods

2.1. Data source and collection

To facilitate a coherent review, we use systematic searches and formal summaries of the literature to integrate major studies in the area. ISI Web of Science (WOS) is used as the data source to collect relevant scholarly work. WOS is the world's leading citation database with a multidisciplinary coverage of over 10,000 high impact journals in science and social science as well as proceedings of over 120,000 international conferences. Specifically, we select the databases within WOS including Science Citation Index Expanded (SCIE), Social Sciences Citation Index (SSCI), Conference Proceedings Index-Science (CPI-S), and Conference Proceedings Index-Social Science and Humanities (CPI-SSH).

We start with an exhaustive search in the databases using combinations of sustainability related keywords (i.e., sustainability, sustainable, green, and social responsibility) with the term of DEA or data envelopment analysis in the fields of title, abstract, author keyword, or Keywords Plus®. A sample of 475 articles is retrieved from the databases after the initial search. We then read through these papers' abstracts to assess whether they dealt with DEA applications in sustainability. When we are unsure, we download and read the full publications. Non-DEA or non-sustainability papers are manually examined and excluded from the dataset. In the course of manual checking and screening, we find out that some papers, although listed as DEA or sustainability in the Keywords Plus® field, contain limited contents about DEA or sustainability For these cases, we conduct a partition analysis on the citation network to find out the outliers and then remove them from the dataset. After the manual checking and screening, the final sample consists of 320 articles published from 1996 to March 2016 in various subject areas including corporate sustainability assessment, sustainability composite indicators construction, sustainability performance analysis, and regional sustainability development assessment. Out of the 320 articles, 120 were published in journals with a 2015 ABS (Chartered Association of Business Schools) journal ranking of 3 or above. We then developed a detailed summary for each article in the final sample.

2.2. Citation-based review methods

To further probe the origins and current state of the literature, we employ two citation-based methods: the main path analysis and Kamada–Kawai algorithm. We believe these citation-based methods can complement to the traditional qualitative review methods by bringing a level of objectivity and quantification. In recent years, the citation-based methods have been applied increasingly across a variety of research fields such as literature research (Liu, Lu, Lu, & Lin, 2013), journal evaluation (Garfield, 1972), and scholar assessment (Schildt, Zahra, & Sillanpää, 2006).

The main path analysis, introduced by Hummon and Doreian (1989), is a well-known method that traces the main knowledge flow in a scientific discipline through citation data. This network-based method treats scientific publications as nodes of a network. Then, citation information is used to establish links among nodes, and a link's direction points from the cited document to the citing one. The first step in finding the main path is to identify the importance of each citation link in the network, which can be measured by counting the times a citation link has been traversed. In this study, we choose to use the search path count (SPC) recommended by Batagelj (2003) to do the counts. The SPC for each link is defined as the total number of times a link is