Measuring and evaluating safety management system effectiveness using Data Envelopment Analysis

Alan J. Stolzer*, Mark A. Friend, Dothang Truong, William A. Tuccio, Marisa Aguiar

Department of Doctoral Studies, Embry-Riddle Aeronautical University, 600 S. Clyde Morris Blvd., Daytona Beach, FL 32114, USA

ARTICLE INFO

ABSTRACT

In January 2015, the FAA released a congressionally mandated, final rule on Safety Management Systems (SMS) for Part 121 air carriers. Since organizations must implement SMS, often at considerable expense, a sound, valid means of measuring SMS effectiveness must be established. The purpose of this research was to develop a model to measure and test SMS effectiveness. It was constructed through a systematic series of steps, ensuring accomplishment of research objectives. While preliminary, this research demonstrates that Data Envelopment Analysis (DEA) models can be produced to help organizations measure the effectiveness of their SMS and determine how to improve SMS-related performance.

1. Introduction

In response to requirements established by the International Civil Aviation Organization (ICAO) in January 2015, the Federal Aviation Administration (FAA) enacted rules for certain certificate holders, specifically CFR14 Part 121 air carriers, to implement a Safety Management System (SMS) for flight-related aspects of the business. An SMS is an organization-wide approach to managing safety and assuring risk-control effectiveness (FAA, 2015); the overall goal of SMS is to ensure safe operation of aircraft through effective management of safety risk (ICAO, 2013). The system is designed to continuously improve safety by identifying hazards, collecting and analyzing data, and assessing risks. Organizations must be able to determine whether the SMS in place is effective and working properly. In order to meet this objective, companies need a reliable and valid instrument for measuring SMS effectiveness. In many cases, they have no formal method of assessment at all. During the interviews with aviation organizations, researchers determined organizations often evaluate the effectiveness of their SMS by reviewing data and considering trends. Trends and data are typically reviewed at regular intervals by safety departments and mid-level management (e.g., monthly) and less frequently by members of top management (e.g., quarterly). In some cases, organizations establish acceptable levels of accidents and incidents and, only when problems in a given area exceed those levels, do they intervene. As long as they are not observing excessive numbers, they operate under the assumption that the SMS is performing effectively.

Although the specific components of SMS are clearly defined and phases outlined, there is no clear path for evaluating its effectiveness.

Literature review shows that some efforts have been made toward building evaluation tools, but close examination of those tools always reveals short-comings. Thomas (2012) reviewed 2009 articles and found only 18 studies used self-report metrics about perceptions of safety within the organization to examine the effectiveness of an SMS. He found a general lack of consistency in the relationship of SMS elements to safety. Clear indicators of success also appear to be lacking. For example, Transport Canada's (2005) assessment tool simply evaluates an enterprise on whether it has a policy in place and a few other criteria that appear to do little to help evaluate its effectiveness. There are few performance criteria in any of the current evaluation tools; therefore, much is left to the judgment of the person performing the assessment. Finally, the literature review indicates existing instruments for measuring SMS effectiveness have not been validated.

By integrating information found in the literature with that supplied by aviation organizations through interviews and surveys, the researchers were able to develop a Data Envelopment Analysis (DEA) model for determining SMS effectiveness in an organization. DEA is a quantitative, analytical programming technique originally developed by Charnes et al. (1978) and is used for measuring and evaluating the performance of Decision-Making Units (DMUs). Over the past 30 years, DEA modeling application has been used to evaluate performance within hospitals, universities, US Air Force, and various other business and legal firms (Cooper et al., 2011). DEA modeling also requires very few assumptions allowing for use in cases that have been resistant to other models due to over complexity or many unknown factors (Cooper et al., 2011). The following objectives guided this research:

- Discover methods existing in other industries to inform and enhance the development of an SMS effectiveness model and determine whether the methods could be adapted for purposes of this project.
- Determine the essential elements to be evaluated in a common model scalable to the size of the organization.
- Develop, test, and validate an instrument for measuring SMS effectiveness.
- Develop and test the Data Envelopment Analysis (DEA) model with several aviation organizations to calculate their efficiency score, identify ineffective organizations, and determine areas of improvement.

Responses requested in the survey were based on the research and interviews conducted prior to the initial pilot study. The research was designed to develop a model for evaluating the effectiveness of a Safety Management System, as defined by the FAA. The tool constructed was intended strictly for use by aviation service providers, such as airlines and airports. The researchers attempted to address all critical features of an SMS and all responses were considered to have appropriately addressed the scope and depth of the SMS within the organization of each participant. Use of this tool by manufacturers in the aviation industry and other entities, not directly related to passenger and cargo transportation, were not included in this study. Respondents to the survey were informed that all answers would remain anonymous and no respondent identities or information disclosed could be tied to a particular respondent nor their organization; therefore, the assumption was made that all answers were truthful.

Following the introduction (Section 1), the paper reviews the literature (Section 2) and discusses the methodology in terms of the steps followed and the data analysis technique (Section 3). It then addresses the results of the interviews (4.1), survey demographics (4.2), instrument reliability and validity test (4.3), and the DEA models (4.4). Conclusions and recommendations for future research are addressed in the final section (5.0).

2. Literature review

A comprehensive literature review was conducted using databases within the Embry-Riddle Aeronautical University Hunt Library and various online sources. SMS effectiveness studies pertaining to various industries, in particular transportation industries, were identified and analyzed.

According to ICAO (2013), an *SMS* is a system to ensure safe operation of aircraft through effective management of safety risk. The system is designed to continuously improve safety by identifying hazards, collecting and analyzing data, and continuously assessing risks. The primary goal of an SMS is to proactively contain or mitigate risks to preclude resulting accidents and incidents. Arendt and Adamski (2011, p. 1) state:

"Sound safety planning, including hazard identification, risk management and safety assurance must be based on a thorough understanding of the processes and activities of people in the system, and the other components of the systems and environments in which they work."

Measuring and controlling performance is an essential part of the process. Setting goals, identifying activities to reach those goals, and improving performance are all subcomponents of that process. This requires measuring performance against pre-established performance-level expectations and implementing changes to adjust to acceptable levels. Safety performance is measured with the same tools and techniques utilized in quality management. Standards against which they are measured are global in nature (Janicak, 2009).

The safety program should be scaled to the size and complexity of the operation and incorporate a mechanism for maintaining and evaluating its effectiveness based on the four components of an SMS (FAA, 2015):

1. *Safety policy and objectives*. The safety policy establishes senior management's commitment to continually improve safety. It defines the methods, processes, and organizational structure needed to meet safety goals.
2. *Safety risk management (SRM)*. SRM determines the need for and adequacy of new or revised risk controls based on the assessment of acceptable risk.
3. *Safety Assurance*. Safety assurance evaluates the continued effectiveness of implemented risk control strategies and supports the identification of new hazards.
4. *Safety Promotion*. Safety promotion includes training, communication, and other actions to create a positive safety culture within all levels of the workforce.

Successful implementation of an SMS hinges on the active participation of every employee in fulfilling their designated roles (Chen and Chen, 2014). An SMS emphasizes the integration of the entire organization serving as one team to proactively manage safety (Chen and Chen, 2014). Specifically, it is important for management to understand that an SMS is accomplished through implementation in a series of phases (ICAO, 2013). This four-phase implementation embodies the aspects of proactive data collection, information analysis, hazard identification, risk management, auditing, training, and reactive incident and accident analysis (Chen and Chen, 2012). Plans for implementation should always consider that similar approaches encourage a flat organizational structure (Yu and Hunt, 2004). Enterprises should be open to change in culture and shifts in management philosophies (Yu and Hunt, 2004). Reports suggest there are five indications of an effective safety culture that must be met, including organizational commitment, management involvement, employee empowerment, reward systems, and reporting systems (Remawi et al., 2011). This requires full commitment and participation by all to include significant time commitments in meetings and ongoing follow-up activities (Yu and Hunt, 2004). According to Yu and Hunt (2004), there must also be clarity of roles and expectations. Executives must embrace a more democratic style of management, yet make it clear that although individual ideas will not always be accepted and implemented, they will be considered. Based on this concept, executive management and policy makers are obligated to demonstrate their appreciation of safety and commitment to SMS practices (Chen and Chen, 2012). ICAO (2013) outlines a four-phase implementation of SMS to include the following steps and timelines:

Phase 1: During this phase, basic planning and assigning of responsibilities occurs. An implementation team and plan are established, and a gap analysis is performed. Key safety personnel are appointed, training and education planned, and a strong safety communication system is put into place. Phase 1 is expected to take approximately 12 months.

Phase 2: This phase consists of implementing essential safety management processes while correcting potential deficiencies in safety management processes. Safety policies will be developed and communicated, accountabilities established, the emergency response plan (ERP) coordinated, and an SMS documentation system set up and made operational. Phase 2 is also expected to take approximately 12 months.

Phase 3: The objective of Phase 3 is to establish safety risk management processes. By the end of this phase, the organization will be ready to collect safety data and perform analysis based on information found. This phase involves managing change and developing processes and documentation for continuous improvement. Phase 3 is expected to take approximately 18 months.

Phase 4: This phase involves mature implementation of safety risk management and safety assurance. All of the above elements,