Transportation Research Part A 92 (2016) 349–364

Contents lists available at ScienceDirect

Transportation Research Part A

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

Benchmarking airports with specific safety performance measures

Leonard MacLean [a,*], Alex Richman [b], Stuart MacLean [c]

[a] Rowe School of Business, Dalhousie University, Halifax B3H 3J5, Canada
[b] AlgoPlus Consulting Ltd., Halifax, Canada
[c] Dept. of Computer Science, York University, Toronto, Canada

ARTICLE INFO

Article history:
Available online 25 June 2016

Keywords:
Aviation safety
Performance measures
Risk priority
Benchmarking

ABSTRACT

The purpose of this paper is to develop safety performance measures and test the measures on data for air traffic management failure events. Failure events are classified by the severity of the consequence of occurrence, resulting in the rate of occurrence in severity categories. The safety measures are standard statistics calculated from this "distribution" for comparison of airport operations by stochastic ordering. For comparisons a benchmark is developed from the aggregation of failure data on a set of comparable airports. Airport performance is then compared with the benchmark using the defined safety measures. The benchmark comparison was implemented with failure data for major airports in Canada from 2005 to 2012. The results show a number of patterns and anomalies and some airports perform poorly in comparison to a class of similar operations. We conclude by suggesting benchmarking safety measures as a natural addition to the information system on aviation safety compiled by a national regulatory body to unravel anomalies such as implementation problems of a safety management system.

1. Introduction

The management of incoming and outgoing traffic of aircraft at airports is a major function of air traffic control. Traffic controllers follow procedures and guidelines established by past practice, industry regulations and company policies. The procedures seek to ensure flight safety while enabling efficient operations. With the competition and financial pressure on airports and carriers, there is an increasing need for efficiency, and this could be at the expense of safety.

Benchmarking is a standard method of assessing performance by comparison of processes with similar operations at a target or ideal site, to identify gaps on a variety of measures. For airports the Air Transport Research Society produces an annual Global Airport Benchmarking Report (Oum et al., 2014), which compares a wide selection of international airports based on Financial, traffic and capacity data. Airport safety measures are not included in that analysis.

A key component of comprehensive benchmarking is the development of agreed upon measures of performance. The Performance review Commission of the European Organization for Safety of Air Navigation has reported on safety benchmarking. However, there is not a common agreement on key indicators of safety and the effort has had limited success. The Transportation Research Board has a detailed guide on performance measures (Hazel et al., 2011), and key safety measures are included. For example, aircraft accidents and incidents during take-off and landing are identified.

* Corresponding author.
E-mail address: l.c.maclean@dal.ca (L. MacLean).

http://dx.doi.org/10.1016/j.tra.2016.06.016

Safety risk is usually managed by a Safety Management System (SMS), which considers multiple dimensions including safety culture. Bala et al. (2013) describe some of the research on the many aspects of aviation safety. Multi-criteria decision making has been used to create safety indices which bring together the many attributes of safety. Chang and Yeh (2004) define a safety index as a weighted combination of scores on criteria, with the weights determined by the Analytic Hierarchy Process (AHP: Saaty, 1980). Liou et al. (2007) develop weights for an index using the Analytic Network Process (ANP: Saaty 1996), where the criteria can be dependent. Chang et al. (2015) use the ANP approach to evaluate airport safety management in Taiwan.

The systematic approach with SMS should lead to better outcomes in terms of accidents and other incidents at airports. The accident data are a hard measure of safety and have been used to develop an index (Li et al., 2014). However, safety is not only accident rates, and often the rate is low and insensitive. Cardoso et al. (2008) consider accidents and other incidents such as airspace violations in estimating performance indicators. The indicators were developed by airport operators.

When incidents or accidents occur at airports a record is generated. Transport Canada collects aviation safety occurrence information through CADORS – Civil Aviation Daily Occurrence Reporting System. This real time data base is rich with information relating to airport operations. It classifies occurrences by the criteria deemed relevant by Transport Canada. The information in CADORS reports is sufficient to assign a consequence weight to occurrences. The events at an airport, therefore, are a marked point process. With this perspective the comparison of airports requires an ordering of random processes. Stochastic ordering or dominance (Levy, 1998) is well developed in probability and is a foundation for measures which are suitable for a safety analysis of airports and the benchmarking of individual airports over time against safety performance standards.

Transport Canada is responsible for both the regulation of aviation and the operation of air traffic services. The safety performance measures provide a framework for establishing standards and monitoring individual airports operational safety against benchmarks implied by performance standards. The feedback to airports in the form of comparative trajectories on safety measures should lead to improved safety management procedures.

The objective of this paper is to present safety measures for airports, which are based on principles of stochastic dominance and are easily computable from the reports which are completed for The Transportation Safety Board of Canada, which is the official source of aviation accident and incident data in Canada. A key component in the analysis is the severity scoring of an incident/accident using guidelines from the US department of Defense (2000). In Section 2 of this paper the record of operations at an airport is considered from a marked point process perspective. The air traffic failures process is defined and the key characteristics or components of that stochastic process are identified. Special attention is given to the measurement of extreme events with an exceedance measure. The components are combined into safety indices and the validity of the indices based on concepts of stochastic dominance is presented. The safety methods are applied to data on major Canadian airports in Section 3. Benchmarks for airport safety performance are developed by grouping airports with similar characteristics (Aerts and Houwing, 2015).

2. Air traffic failure process

The operation of airports in the management of air traffic is complex, with many interrelated systems and procedures. The importance of effective performance has resulted in various performance indicators being proposed by national and international organizations (Hazel et al., 2011). Included in those indicators are measures of safe operations. There is some variation in the lists of measures of airport safety, but the measures which relate directly to the successful movement of aircraft to and from locations are the key ones.

Let $m(t)$ be the number of movements to time t at the airport(s) of interest. Movement times are scheduled, so it will be assumed that $m(t)$ is a smooth function of t. For simplicity, let $m(t) = a_1 + a_2 t$, so that accumulated use (movements) are linear in time and the rate of movements is constant.

If a failure event occurs during a movement at time t, there is a consequence associated with that failure. The severity of failures can be wide ranging and the associated risk factors and consequences of failure depend on the airport. None the less, it is common to classify a failure using a standard set of ordered classes ranging from "no consequences" to "extreme consequences", where extreme would involve death/destruction. The allocation of a failure event to a severity class is a judgement, using established criteria. So the data on severity for a failure event records a severity score, but that is typically the class number on an ordinal classification scheme. The resulting classification scheme has a "description" of the defining characteristics/consequences for each class. The event will have effects in multiple dimensions: people, equipment, schedules. Based on the fact that the classes are ordered, it is assumed that there is a measure which maps the vector of effects into the real number line. In the simplest interpretation each dimension has a cost multiplier and the total event effect is the sum of the cost in each dimension. So the range of total effects is an interval on the real line, possibly $(0, \infty)$. Then the half-line is partitioned into intervals associated with the severity classes.

Let the severity of a failure at time s be $X(s)$, where $X(s) \in R$. Consider, then, for incoming and outgoing traffic (movements) at an airport operating between times 0 and t, the observed failure events: incidents and accidents of varying severity at random times – $X(t_i),\ i \in A(t)$, where $A(t)$ is the set of failure times. Those events provide a record of the air traffic safety at an airport. The accumulation of failures over time is a measure of the growth in defective processes for the safety management system. Consider, then, the accumulated failure severities to time t: