Physica A 463 (2016) 345–355

Contents lists available at ScienceDirect

Physica A

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

Quantifying the risk of extreme aviation accidents

Kumer Pial Das [a,*], Asim Kumer Dey [b]

[a] Lamar University, TX, USA
[b] University of Texas at Dallas, TX, USA

HIGHLIGHTS

- The study focuses on a model involving extreme aviation accidents.
- The fitted generalized Pareto model has been compared with few other possible models.
- Bootstrap resampling technique provides accurate estimate of the uncertainty.
- The return level estimates have been tested for several threshold levels.

ARTICLE INFO

Article history:
Received 27 December 2015
Received in revised form 16 May 2016
Available online 25 July 2016

Keywords:
Aviation accident
Extreme value model
Return level
Uncertainty
Bootstrap sampling
Monte Carlo simulation

ABSTRACT

Air travel is considered a safe means of transportation. But when aviation accidents do occur they often result in fatalities. Fortunately, the most extreme accidents occur rarely. However, 2014 was the deadliest year in the past decade causing 111 plane crashes, and among them worst four crashes cause 298, 239, 162 and 116 deaths. In this study, we want to assess the risk of the catastrophic aviation accidents by studying historical aviation accidents. Applying a generalized Pareto model we predict the maximum fatalities from an aviation accident in future. The fitted model is compared with some of its competitive models. The uncertainty in the inferences are quantified using simulated aviation accident series, generated by bootstrap resampling and Monte Carlo simulations.

1. Introduction

Aviation accidents cause enormous loss of lives and massive monetary costs worldwide. In 2002–2011, there were a total of 250 worldwide fatal accidents, which resulted in 7148 fatalities [1]. Among these 250 fatal accidents, around 20% accidents can be considered as extremes and each of them was responsible for more than 50 fatal injuries. Although there is an overall decreasing trend in the number of fatal accidents [2], there is no such trend in the number of extreme fatal accidents which cause huge fatality. Quantification of large accidents which have far reaching effect, in terms of fatality and monetary loss, would provide objective guidance in long-term planning and response for manufacturers, insurers and re-insurers.

Every year different organizations like Federal Aviation Administration (FAA), International Civil Aviation Organization (ICAO), Civil Aviation Authority (CAA) of UK and leading manufacturers of commercial jetliners such as Boeing, Airbus give valuable reports on aviation accident. These reports mainly focus on descriptive accident statistics such as the rate of worldwide fatal accidents and fatalities by year, nature of flight, aircraft age and weight group [3,4]. Sometimes they make analysis on different causal factors for these accidents [1]. Some research on aviation accident focus on behaviors

* Corresponding author.
E-mail addresses: kumer.das@lamar.edu (K.P. Das), adey@utdallas.edu (A.K. Dey).

http://dx.doi.org/10.1016/j.physa.2016.07.023
0378-4371/

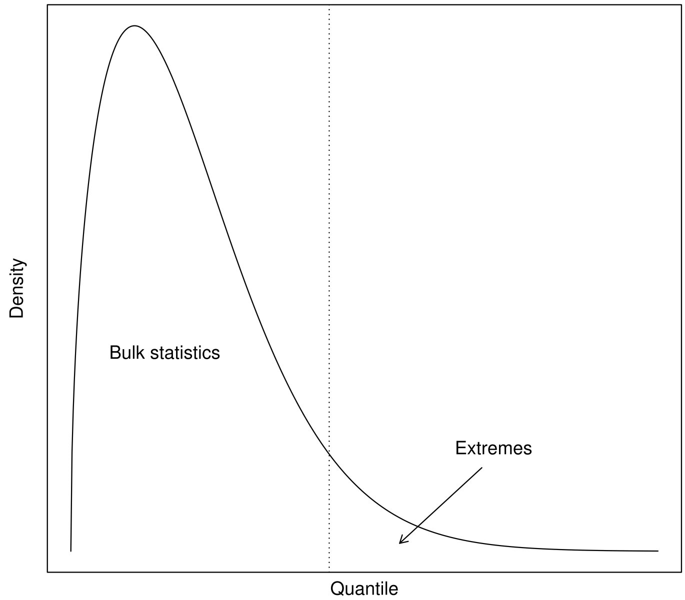

Fig. 1. Extreme value model.

that are associated with Loss-of-Control (LOC) events [5]. Some others calculate the occurrence probabilities of serious incidents using safety critical measures runway overrun, which indicate the actual landing distance [6]. But there is no well known research focusing on the modeling of the number of fatal injuries in individual aviation accidents. Using different tail models (power law, stretched exponential and log-normal), one can assess the likelihood of a large event like serious aviation accidents [7]. The classical extreme value theory, for example, generalized Pareto distribution (GPD) can also be used to model extreme aviation accidents [8].

In this paper we study aviation accidents from 1982 to 2014, their pattern within the period, the number of fatalities from them, etc. A prediction of possible number of fatal injuries for an extreme aviation accident in the future is made using peaks over threshold method (i.e., generalized Pareto distribution approach). We compare the fitted generalized Pareto distribution (GPD) model with other long tail models to select better model for aviation accident using different measure of goodness of fit. We also quantify the uncertainty in the inferences using data generated by bootstrap resampling and Monte Carlo simulation. The overall aim of this paper is to measure the risk of aviation accident in terms of fatal injury. The challenge is in selecting reasonable tail model for this data as well as in measuring uncertainty in the inferences.

Rest of the paper is organized into five sections: Section 2 introduces a brief description of the extreme value models and the generalized Pareto distribution (GPD); Section 3 proposes a GPD model for aviation fatal injury; Section 4 describes different uncertainty measurements in the inference of extreme fatalities through bootstrap resampling and simulated accident series; Section 5 analyzes goodness of fit of the fitted model and also made a comparison among different possible model in this situation; and finally, conclusions appear in Section 6.

2. Extreme value model and generalized Pareto distribution (GPD)

In many statistical applications, the interest is centered on estimating some population characteristics such as average or median of a process based on random samples taken from a population under study. For instance, in regression analysis if data contain outliers then robust estimation procedure is used. A robust estimation procedure is one that dampens the effect of observations that would be highly influential [9]. However, in extreme value analysis, we are not interested in estimating the average rather we want to quantify the behavior of the process at unusually large or small levels. Extreme value theory (EVT) deals with the extreme deviations from the median of probability distributions and seeks to assess, from a given ordered sample of a given random variable, the probability of events that are more extreme than a certain large value. Usual bulk statistics tries to describe main part of a distribution and may ignore outliers. But EVT tries to characterize the tail of the distribution; keeps only the extreme observations (Fig. 1).

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed (i.i.d.) random variables with common distribution function F. Extreme value analysis focuses on the statistical behavior of the maximum value observed, i.e.,

$$M_n = \max\{X_1, X_2, \ldots, X_n\}.$$

In applications, the X_i usually represent values of a process measured on a regular time-scale at time i such as the hourly measurements of sea level, or daily mean temperature so that M_n represents the maximum of the process over n time units