



HAL
open science

Prediction of U.S. General Aviation fatalities from extreme value approach

Abdoulaye Diamoutene, Bernard Kamsu-Foguem, Farid Noureddine, Diakarya Barro

► **To cite this version:**

Abdoulaye Diamoutene, Bernard Kamsu-Foguem, Farid Noureddine, Diakarya Barro. Prediction of U.S. General Aviation fatalities from extreme value approach. *Transportation Research Part A: Policy and Practice*, 2018, 109, pp.65-75. 10.1016/j.tra.2018.01.022 . hal-02135111

HAL Id: hal-02135111

<https://hal.science/hal-02135111>

Submitted on 21 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Open Archive Toulouse Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of some Toulouse researchers and makes it freely available over the web where possible.

This is an author's version published in: <https://oatao.univ-toulouse.fr/22976>

Official URL : <https://doi.org/10.1016/j.tra.2018.01.022>

To cite this version :

Diamoutene, Abdoulaye  and Kamsu-Foguem, Bernard  and Nouredine, Farid  and Barro, Diakarya *Prediction of U.S. General Aviation fatalities from extreme value approach.* (2018) *Transportation Research Part A: Policy and Practice*, 109. 65-75. ISSN 0965-8564

Any correspondence concerning this service should be sent to the repository administrator:

tech-oatao@listes-diff.inp-toulouse.fr

Prediction of U.S. General Aviation fatalities from extreme value approach

Abdoulaye Diamoutene^{a,b,c}, Bernard Kamsu-Foguem^{a,*}, Farid Nouredine^a,
Diakarya Barro^c

^a Université de Toulouse, Laboratoire de Génie de Production (LGP), EA 1905, 47 Avenue d'Azereix, BP 1629, 65016 Tarbes Cedex, France

^b Institut Polytechnique Rural de Formation et de Recherche Appliquée de Katibougou Mali, BP: 06 Koulikoro, Mali

^c Université Ouaga 2, UFR-SEG, 12 BP: 417 Ouagadougou 12, Burkina Faso

ARTICLE INFO

Keywords:

General Aviation accident
Extreme value statistics
Graphical diagnosis
Return level
Uncertainty
Nonparametric bootstrap

ABSTRACT

General Aviation is the main component of the United States civil aviation and the most aviation accidents concern this aviation category. Between early 2015 and May 17, 2016, a total of 1546 general aviation accidents in the United States has left 466 fatalities and 384 injured. Hence, in this study, we investigate the risk of U.S. General Aviation accidents by examining historical U.S. General Aviation accidents. Using the Peak Over Threshold approach and Generalized Pareto Distribution, we predict the number of fatalities resulting in extreme GA accidents in the future operations. We use a graphical method and intensive parameters estimates to obtain the optimal range of the threshold. In order to assess the uncertainty in the inference and the accuracy of the results, we use the nonparametric bootstrap approach.

1. Introduction

General Aviation (GA) is the term for all civil aviation operations other than scheduled air service and non scheduled air transport operations for remuneration or hire, see [Crane \(1997\)](#) and National Transportation Statistics [NTS \(2016\)](#). GA flights range from gliders and powered parachutes to corporate business jet flights, see National Transportation Safety Board [NTSB \(2011\)](#) and Civil Aviation Authority [CAA \(2013\)](#). The majority of the world's air traffic falls into this category and most of the world's airports serve general aviation exclusively, see [Crane \(1997\)](#). The term General Aviation is a catch all phrase for all aviation activities that do not fall under commercial aviation, major cargo or military operations, and covers a broad spectrum of aviation activity uses.

Between early 2015 and May 17, 2016, a total of 1546 general aviation accidents in the United States has left 466 fatalities and 384 injured, according to data from the NTSB.

In 2011, GA aircrafts were involved in 95% of all aviation accidents and 94% of fatal aviation accidents; accidents involving GA aircrafts accounted for 92% of all U.S. civil aviation fatalities, see [NTSB \(2011\)](#); it is an important component of civil aviation in the United States, see [Shetty and Hansman \(2012\)](#).

GA covers a large range of activities, both commercial and non commercial, including flying clubs, flight training, agricultural aviation, light aircraft manufacturing and maintenance. GA is particularly popular in North America, with over 6,300 airports available for public use by pilots of GA aircraft in the U.S., see [NTSB \(2011\)](#) and [NTS \(2016\)](#).

In 1990 2014, there is in United States GA a total of 8402 fatal accidents which resulted in about 14984 fatalities according to data from the National Transportation Safety Board (NTSB). Among these 8402 fatal accidents, around 26 accidents can be

* Corresponding author.

E-mail address: bernard.kamsu-foguem@enit.fr (B. Kamsu-Foguem).

considered as extreme and each of them was responsible for more than 39 fatalities. Moreover, in extreme value approach, an extreme event is an event which has a low probability, but implies serious consequences. In our context, an extreme accident is an accident that has caused over 39 fatalities. Although there is an overall decreasing trend in the annual number of fatalities, see [NTS \(2016\)](#) and [Shetty and Hansman \(2012\)](#), there is no such trend in the number of extreme fatal accidents which cause enormous fatalities. Prediction and risk management in terms of extreme fatality are more important in insurers, reinsurers and aeronautics industry like Boeing and Airbus, see [Boeing \(2016\)](#) and [Airbus \(2016\)](#). Moreover, based on historical data, the return level of fatalities could help management agency of GA to perform aircraft operations, conditions based maintenance, predictive maintenance and flying conditions.

The safety of civil aviation in the United States is regulated by the Federal Aviation Administration (FAA), for more details see [NTSB \(2011\)](#). The FAA distinguishes between commercial operations and GA operations, see [NTSB \(2011\)](#). There are other organizations like International Civil Aviation Organization (ICAO), Civil Aviation Authority (CAA) of the United Kingdom (UK) and leading manufacturers of commercial jetliners such as Boeing, Airbus give valuable reports on civil aviation accident. These reports mainly focus on descriptive accident statistics such as the rate of fatal accidents, fatalities by year, nature of flight, type of aircraft, transport category, flight hours, for more details, see [NTSB \(2011\)](#) and [NTS \(2016\)](#).

In literature, some works on aviation accidents focus on the cost of GA accidents in U.S., see [Sobieralski \(2012\)](#); [Knetch \(2015\)](#) focuses on the prediction of the accident rate due to pilot total flight hours; [Fultz and Ashley \(2016\)](#), study extreme weather impact in GA accidents in the United States; [Farah and Azevedo \(1952\)](#) focus on safety analysis of passing maneuvers using extreme value theory. [Sha et al. \(2017\)](#) focus on the use of the extreme value theory for forecasting spare parts stocks on the basis of repair data set information, see [Romeijnders et al. \(2012\)](#). [Dey and Das \(2016\)](#) propose an approach to predict extreme aviation fatal injuries. But in literature, it is rare to see research concerning the modeling of the number of fatalities in individual U.S. GA accidents; using extreme value models, notably Peaks Over Threshold (POT) approach. From extreme value statistics, we can model the return level of fatalities due to extreme GA accidents. We distinguish two approaches namely, block maxima approach and Peak Over Threshold (POT) approach, see [Coles \(2001\)](#). The block maxima approach is wasteful of data as only one data point in each block is taken. The second highest value in one block may be larger than the highest of another block and this is generally not accounted for, see [Coles \(2001\)](#) and [Castillo et al. \(2005\)](#). Peak Over Threshold (POT) approach avoids this drawback, by using the Generalized Pareto Distribution (GPD). Indeed, due to advances in extreme value theory, the GPD emerged as a natural family for modeling exceedances over a high threshold. The POT approach has shown its importance and success in a number of statistical analysis problems related to different areas, see [Schmidt et al. \(2014\)](#) and [Castillo et al. \(2005\)](#). However, despite the sound theoretical basis and wide applicability, the fitting of this distribution in practice is not a trivial exercise.

In this paper we focus on U.S. GA fatal accidents from 1990 to 2014 and their resulting fatalities number. A prediction of the possible number of fatalities for an extreme GA accident in the future operation is made using POT approach. We use three estimation methods to estimate the parameters and select the method which standard error is less than in the other two methods. The two main factors which affect the accuracy of the estimations of the return values are the choice of the threshold on one hand and the choice of the parameter estimator on the other hand. An extensive discussion of the various parameter estimation methods has been given by [Bermudez and Kotz \(2010\)](#). [Scarrot and Macdonald \(2012\)](#) reviewed the threshold estimation methods; hence, we use graphical method and intensive estimate of GPD parameters to obtain a reasonable range of the threshold. We also quantify the uncertainty in the inference using nonparametric bootstrap approach, see [Carpenter and Bithel \(2000\)](#) and [Yongcheng \(2008\)](#).

This paper is structured as follows: following this introduction, Section 2 provides background on extreme values statistics, notably both block maxima approach and POT approach. Section 3 exposes GA accident modeling with POT approach. Section 4 presents different uncertainty measurements in the inference of U.S. GA extreme fatalities and finally Section 5 concludes and discusses future challenges.

2. Extreme value theory

Most statistical methods are concerned primarily with what goes on in the center of a statistical distribution, and do not pay particular attention to the tails of a distribution, or in other words, the most extreme values at either high or low end. However, in extreme value analysis, we are not interested in estimating the average we rather want to quantify the behavior of the process at unusually large or small levels, see [Coles \(2001\)](#). 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, see [Coles \(2001\)](#) and [Castillo et al. \(2005\)](#).

Broadly speaking, there are two principal kinds of models for extreme values. The first group of model, called Generalized Extreme Value theory (GEV) consists of block maxima approach; these are models for the largest observations collected from large samples of identically distributed observations. A second group of EVT models contains the Peak Over Threshold (POT) models, which model all large observations that exceed a high threshold, for more details see [Coles \(2001\)](#) and [Castillo et al. \(2005\)](#). These POT models are generally considered to be most useful for practical applications, due to their more efficient use of the (often limited) data of extreme values. Within the POT class of models, several styles of analysis exist, see [Coles \(2001\)](#) and [Schmidt et al. \(2014\)](#). Among all these parametric models, the Generalized Pareto Distribution (GPD) has the advantage of being conceptually simple and implemented in software R, see [Ribatet \(2012\)](#).

2.1. Block maxima approach

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

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 of observation, see [Coles \(2001\)](#). If n is the number of observations in a year, then M_n corresponds to the annual maximum. Using the fact that X_1, X_2, \dots, X_n are i.i.d random variables

$$\begin{aligned} Pr(M_n \leq x) &= Pr\{X_1 \leq x, \dots, X_n \leq x\} \\ &= Pr\{X_1 \leq x\} \times \dots \times Pr\{X_n \leq x\} \\ &= \{F(x)\}^n. \end{aligned}$$

In practice, we might not know the distribution function F but according to the extremal type, if there exist sequences of constants $\{a_n > 0\}$ and $\{b_n\}$ such that $Pr\left\{\left(\frac{M_n - b_n}{a_n}\right) \leq x\right\} \rightarrow G(x)$ as $n \rightarrow \infty$ with G being a non degenerate distribution function, then G belongs to the following family of models having a distribution function of the form:

$$G(x) = \exp\left\{-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\}, \quad (1)$$

defined on the set $\left\{z: 1 + \xi\left(\frac{x-\mu}{\sigma}\right) > 0\right\}$, where the parameters satisfy $-\infty < \mu < \infty, \sigma > 0$ and $-\infty < \xi < \infty$. This is the generalized extreme value family of distributions. The model has three parameters: a location parameter μ ; a scale parameter σ ; and a shape parameter ξ . The shape parameter ξ governs the tail behavior of the distribution. The sub families defined by $\xi \rightarrow 0, \xi > 0$ and $\xi < 0$ correspond, respectively to Gumbel, Fréchet and Weibull families, for more details see [Coles \(2001\)](#).

2.2. Peak over threshold approach

Compared to the traditional approaches, the POT method can utilize more information from the data set. In contrast to block maxima approach, it is more practical to analyze the value of random variables that exceed a given threshold value if an entire time series of, say, hourly or daily or yearly observations is available, this method is commonly known as POT approach, see [Far and Wahab \(2016\)](#) and [Castillo and Daoudi \(2009\)](#). Hence, the POT approach provides a more rational selection of events fulfilling the criteria of being "extreme". In this approach observations that exceed a given threshold μ are called exceedances over a threshold. POT method has popularly been used to estimate return levels of significant wave height, see [Mackay et al. \(2014\)](#), hurricane damage in [Daspit and Das \(2012\)](#), [Schmidt et al. \(2014\)](#) in the study of traffic load effects on bridges and earthquakes modeling in [Edwards and Das \(2016\)](#). [Pickands \(1975\)](#) proves that for large enough threshold μ , observations x , provided $x > \mu$, approximatively follow a Generalized Pareto Distribution (GPD) with distribution function:

$$F(x) = \begin{cases} 1 - \left(1 + \frac{\xi(x-\mu)}{\sigma}\right)^{-\frac{1}{\xi}}, \xi \neq 0, \sigma > 0, \\ 1 - \exp\left(-\frac{x-\mu}{\sigma}\right), \xi = 0, \sigma > 0. \end{cases} \quad (2)$$

The GPD defined in Eq. (2) reduces to a 2 parameter GPD for $\mu = 0$ and for most of the practical purposes the 2 parameter GPD seems more appropriate than the 3 parameter GPD as described in [Far and Wahab \(2016\)](#), so in our study, a 2 parameter GPD is considered. For different values of the shape parameter, GPD provides few interesting distributions. For $\xi < 0$, the distribution has a heavy Pareto type upper tail. For $\xi = 0$, the GPD provides the exponential distribution with mean σ . For $\xi = 0.5$ and $\xi = 1$ the distribution is triangular and uniform respectively. When $\xi \leq \frac{-1}{2}$, $Var(X) = \infty$, and the r th central moment exists if and only if $\xi \geq \frac{-1}{r}$. The GPD is stable with respect to excesses over threshold operations. In other words, if the random variable X has a generalized Pareto with $GPD(\xi, \sigma)$, then the conditional distribution of $(X - \mu)$ subject to $X \geq \mu$ is also generalized Pareto with $GPD(\xi, \sigma + \xi\mu)$ and so the new GPD retains the same shape parameter value ξ and this property is known as the threshold stability property as shown in [Mackay et al. \(2014\)](#).

2.3. Graphical diagnosis for threshold choice

The threshold stability property of GPD means that if the GPD is a valid model for excesses over threshold μ_0 , then it is valid for excesses over all thresholds μ_0 , the expected value of our threshold excesses, conditional on being greater than the threshold μ_0 , is

$$E[X - \mu | X > \mu] = \frac{\sigma_{\mu_0} + \xi\mu}{1 - \xi}.$$

Thus, for all $\mu > \mu_0, E[X - \mu | X > \mu]$, is a linear function of μ . Furthermore, $E[X - \mu | X > \mu]$ is simply the mean of the excesses of threshold μ , for which the simple mean of the threshold excesses of μ provides an estimate. This leads to the mean residual life plot (mrlplot) or mean excess plot (meplot), a graphical procedure for identifying a suitably high threshold for modeling extremes via the

GPD. In this plot, for a range of candidate values for μ , we identify the corresponding mean threshold excess; we then plot this mean threshold excess against μ , and look for value μ_0 above which we can see linearity in the plot, for more details see Ribatet (2012) and Scarrot and Macdonald (2012).

2.3.1. Mean excess plot and mean residual life plot

The observations exceeding a given threshold μ in the tail of the distribution approximately follow GPD. However, a crucial step in this analysis is to select an appropriate threshold μ . On one hand, if μ is too small, then the GPD will not fit the tail distribution and the estimates of parameters (ξ, σ) may be biased. On the other hand, if μ is too large, the GPD may fit well, but fewer observations will be available to estimate the parameters and the estimates will suffer from increased variance. We want to select the smallest μ in mrplot and meplot (Fig. 3) above which the mrplot and meplot will be approximately linear.

2.3.2. Parameter stability plot

A common diagnosis is the threshold stability plot where estimated shape parameters (including pointwise uncertainty intervals) are plotted against a range of possible thresholds. The threshold is chosen as the lowest possible value such that the shape parameter is stable for all higher values, once sample uncertainty is taken into account. A similar plot of modified scale parameter estimated is occasionally considered, although the correlation between scale and shape parameter is often not obvious.

In practice, threshold μ should be chosen where the shape and modified scale parameters remain constant in parameter stability plot after taking the sampling variability into account.

2.4. Return period and return level estimation

In extreme value statistics it is common to use the notion of return period and return level to interpret and explain the information about the likelihood of extreme events such as floods, earthquakes, hurricanes, aviation accidents, etc. An estimate of the t observations return level x_t that is exceeded on average once every t observations is given by

$$\hat{x}_t = \begin{cases} \mu + \frac{\hat{\sigma}}{\xi} [(t\hat{\lambda}_\mu)^\xi - 1] \text{ for } \xi \neq 0 \\ \mu + \hat{\sigma} \log(t\hat{\lambda}_\mu) \text{ for } \xi = 0, \end{cases} \quad (3)$$

provided that t is sufficiently large to ensure that $\hat{x}_t > \mu$.

An estimate of λ_μ , the probability of an individual observation exceeding the threshold μ , is also needed. The natural estimator of λ_μ is $\hat{\lambda}_\mu = \frac{k}{n}$, the sample proportion of points exceeding μ . Since the number of exceedances of μ follows the binomial $Bin(n, \lambda_\mu)$ distribution, $\hat{\lambda}_\mu$ is also the maximum likelihood estimate of λ_μ , see Coles (2001).

By construction, x_t is the t observation return level; however, it is often more convenient to give return levels on an annual scale, so that the T year return level is the level expected to be exceeded once every T years. If there are n_y observations per year, this corresponds to the t observation return level with $t = T \times n_y$. Hence, an estimate of the T - year return level x_T is defined by

$$\hat{x}_T = \begin{cases} \mu + \frac{\hat{\sigma}}{\xi} [(T \cdot n_y \hat{\lambda}_\mu)^\xi - 1] \text{ for } \xi \neq 0 \\ \mu + \hat{\sigma} \log(T \cdot n_y \cdot \hat{\lambda}_\mu) \text{ for } \xi = 0, \end{cases} \quad (4)$$

provided that t is sufficiently large to ensure that $\hat{x}_T > \mu$.

Estimation of return levels \hat{x}_T requires the substitution of the estimates of the different parameters and likelihood estimate of λ_μ .

In the remainder of this document, the return period T corresponds to the period of prediction in terms of years and the return level x_T is the possible prediction of fatalities according to model parameters.

3. Modeling of U.S. General Aviation fatalities

3.1. Data presentation

The purpose of this study is to develop a predictive model which will provide estimates of potential fatalities for an extreme aviation accident in the future of U.S. GA. Data are obtained from U.S. General Aviation safety data which is a division of the National Transportation Safety Board (NTSB). The NTSB aviation accident database contains information within the United States, its territories and possessions and in international waters. The data consist of 8402 U.S. GA fatal accidents from 1990 to 2014 which caused in total about 14984 fatalities. U.S. registered civil aircrafts are not operated under 14 CFR 121 or 14 CFR 135. Accidents on foreign soil and in foreign waters are excluded. Suicide, sabotage, and stolen/unauthorized cases are included in total accidents fatalities.

Fig. 1 shows the evolution of U.S. General Aviation fatalities from 1990 to 2014. Generally speaking, we observe a diminution of fatalities over years. However, Fig. 1 does not exhibit an individual extreme fatality.

Fig. 2 shows the individual assignment of the data; we observe that most fatalities is less than 39.

Figs. 1 and 2 show that there is a nuance between the declination trend of annual fatalities and fatalities resulting in individual extreme GA accident.

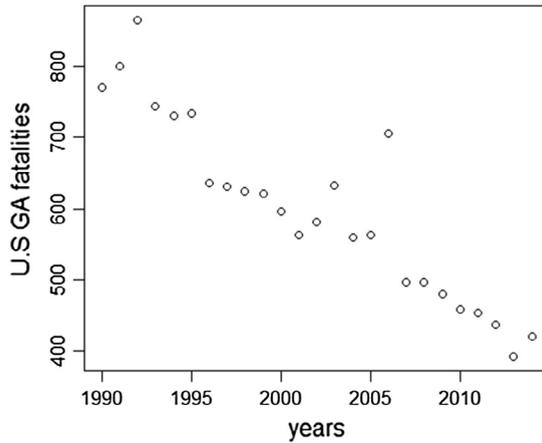


Fig. 1. Annual fatalities.

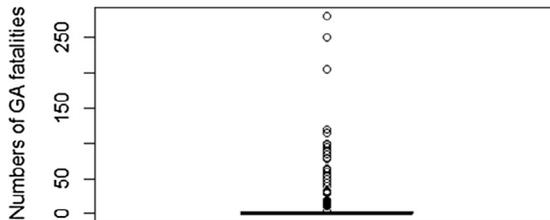


Fig. 2. Boxplot of U.S. GA data.

3.2. Application of graphical approach in the choice of threshold

In this subsection, we use the theoretical approach presented in Section 2.3. We would selected the smallest μ in meplot and mrlplot (Fig. 3) above which the meplot and mrlplot will be approximately linear.

Fig. 3 shows that the reasonable threshold should be in the range (40 70).

According to the Fig. 3, parameter stability plots in Fig. 4 also suggest selecting the threshold in the range (40 50).

3.3. Fitting the GPD parameter and return levels for different threshold values

Once the threshold value has been specified, estimating the shape and scale parameters would be the next step of developing models. The importance of parameter estimations in POT method cannot be underestimated as they may create errors in estimating the high quantiles. There are numerous parameter estimation methods available in the literature to fit the GPD parameters to the set of threshold excesses such as maximum likelihood estimate (MLE), Method of Moment (MOM) and Probability Weighted Moment (PWM), least squares error method, etc. Dupuis and Tsao (1998) proposed a hybrid estimators of two well known estimators: Method of Moment (MOM) and Probability Weight Moment (PWM); the methods were derived by incorporating a simple auxiliary limitation of the estimates. Bermudez and Kotz (2010) carried out an extensive study on several types of methods for estimating the GPD parameters; the authors argued that the success of GPD on a set data depends substantially on the process of parameter estimates, and

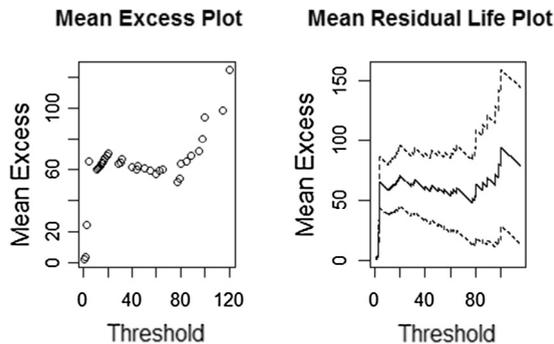


Fig. 3. Mean excess plot and mean residual life plot.

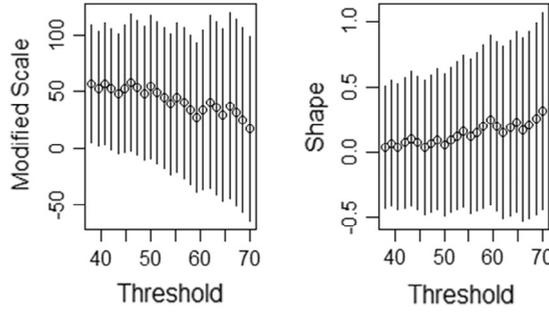


Fig. 4. Parameter stability plot.

(MLE) estimators has been the most popular method among other estimators. However, the (MLE) estimators exist only for the shape parameter $\xi \leq 1$ because for $\xi > 1$ the log likelihood becomes infinite. In this study, the estimation of GPD parameters and adequate threshold choice are performed using the maximum likelihood, Method of Moment and Probability Weighted Moment (PWM).

For maximum likelihood estimate, the GPD log likelihood function for $\xi \neq 0$ can be derived in the usual way as follows:

$$l(\xi, \sigma; y) = -ne \cdot \log \sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{ne} \log \left(1 + \frac{\xi y_i}{\sigma}\right), \quad (5)$$

where $y = (y_1, y_2, \dots, y_{ne})$ are the set of exceedances above threshold μ . For the case $\xi = 0$, interpreted as $\xi \rightarrow 0$, we have the log likelihood for an exponential distribution with rate $\frac{1}{\sigma}$. In the software R, we could write a function which computes the negative log likelihood for the GPD, and then use the *nme* routine to minimize this with respect to $\theta = (\xi, \sigma)$. Here, we will obtain maximum likelihood estimates of the GPD parameters for our aviation fatalities by using the *fitgpd* function in the package *POT*.

Method of Moment estimates of ξ and σ can be easily obtained by using the mean and variance of the exceedances. The estimates of ξ and σ are given by:

$$\hat{\xi} = \frac{1}{2} \left(\frac{\bar{y}^2}{s^2} - 1 \right) \text{ and } \hat{\sigma} = \frac{1}{2} \bar{y} \left(\frac{\bar{y}^2}{s^2} + 1 \right). \quad (6)$$

Probability Weighted Moment (PWM) is based on Hosking and Wallis (1987), the general expression for the r th order of PWM of GPD is given by:

$$M_r = \frac{u}{1+r} + \frac{\sigma}{(1+r)(1+r+\xi)}.$$

The estimate of the r th order of PWM is given by:

$$\widehat{M}_r = \sum_{i=1}^k (1-P_i)^2 \frac{y_{i:k}}{k}$$

with plotting position $P_i = \frac{i-0.35}{k}$. The estimates of ξ and σ can be obtained using

$$\hat{\xi} = \frac{\widehat{M}_0}{\widehat{M}_0 - 2\widehat{M}_1} - 2 \text{ and } \hat{\sigma} = \frac{2\widehat{M}_0\widehat{M}_1}{\widehat{M}_0 - 2\widehat{M}_1}, \quad (7)$$

for more details see Dupuis and Tsao (1998).

Table 1
Parameters estimates in MLE for various thresholds.

μ	$\hat{\xi}$	$\hat{\sigma}$	$se(\hat{\xi})$	$se(\hat{\sigma})$
40	0.02545	60.21724	0.2449	19.8103
41	0.04654	57.96778	0.2492	19.2504
42	0.06782	55.75671	0.2537	18.7157
43	0.09026	53.5635	0.2592	18.2129
44	0.05335	57.4497	0.2588	19.7149
45	0.01133	62.11331	0.26	21.61
46	0.0328	59.7771	0.2634	20.9387
47	0.05326	57.59795	0.267	20.354
48	0.07631	55.28838	0.2714	19.7162
49	0.1035	52.7837	0.2772	19.0017
50	0.05656	57.59245	0.278	21.043

Table 2

Parameters estimates in MOM for the various thresholds.

μ	$\hat{\xi}$	$\hat{\sigma}$	$se(\hat{\xi})$	$se(\hat{\sigma})$
40	0.0458	58.7971	0.2325	18.7472
41	0.06042	56.95646	0.2392	18.4507
42	0.0748	55.1594	0.2473	18.2158
43	0.08895	53.40504	0.2568	18.0524
44	0.06594	56.51083	0.2481	18.8886
45	0.03779	60.26450	0.241	20.053
46	0.05244	58.39990	0.2475	19.7094
47	0.06684	56.57887	0.2551	19.4256
48	0.08101	54.80071	0.2641	19.2103
49	0.09495	53.06471	0.2748	19.0751
50	0.06837	56.67434	0.263	20.032

3.3.1. Maximum Likelihood Estimates of GPD parameters for potential thresholds

The parameter values and their standard errors are computed over formula 5 and software R; the results are in the [Table 1](#).

For $\xi > 0$, the distribution has no upper limit, see [Coles \(2001\)](#). The examination of the shape parameter estimates and its standard error show some level of instability in the shape parameter estimation. In the ranges 40–43 and 45–49, the shape parameter increases and the same for the standard errors, for threshold values 45 and 50 we have the isolate values. We also observe that MLE is sensitive to the threshold choice.

3.3.2. Method of Moment for potential thresholds

[Table 2](#) below gives the values and the standard errors of the parameters over formula 6 and software R.

In MOM, the shape parameter is positive showing that the distribution has no upper limit. The examination of the shape parameter estimates and its standard errors show the same results as MLE but the standard error of shape parameter in MOM is less than in MLE.

3.3.3. Method of Probability Weight Moment for potential thresholds

The [Table 3](#) below provides the values and the standard errors of the parameters over formula 7 and software R.

Concerning PWM, we have the same variation with a greater standard error for the estimates of the shape parameter.

All these methods are sensitive to the threshold choice. The examination on the estimates in the three selected methods shows that MOM and MLE are less sensitive to any change in the value of threshold than PWM. In the following, we use MLE and MOM because their standard errors are less high than in PWM.

3.3.4. General Aviation fatalities return level estimates for the potential thresholds

In this subsection, we use the estimates of the parameters from [Tables 1 and 2](#) for the threshold value 40, 45 and 50 to calculate the return level of fatalities for the periods 5, 10 and 25 years. By replacing in Eq. (4), the estimates of parameters and return period, we obtain the [Table 4](#) below.

[Table 4](#) summarizes the T year return levels for different values of T and the two methods used. The examination of the results in [Table 4](#) shows that for each T return period, the estimates of return levels are almost the same for the two different methods (MLE and MOM). However, from threshold 45 to threshold 50, we note that there is a very little change in fatalities number. This means that the optimal range of threshold is range 45–50. For example, with MOM 5 year return level is 139 fatalities for threshold 45, that implies that about 139 fatalities are likely to occur once in every 5 years. Another interpretation for 5 year return level for threshold 45 with 2014 as reference year, is that in 2019 a possible number of fatalities is 139. General Aviation accidents factors such that technology,

Table 3

Parameters estimates in PWM for the different thresholds.

μ	$\hat{\xi}$	$\hat{\sigma}$	$se(\hat{\xi})$	$se(\hat{\sigma})$
40	0.04856	58.62684	0.2509	19.5998
41	0.08023	55.75563	0.2513	18.6905
42	0.11119	52.9478	0.2528	17.8111
43	0.1436	50.2032	0.2554	16.9630
44	0.09243	54.90782	0.2579	18.8842
45	0.02799	60.87868	0.264	21.367
46	0.05947	57.86614	0.2638	20.3908
47	0.09096	55.11657	0.2646	19.4454
48	0.1224	52.33	0.2665	18.5319
49	0.1539	49.6064	0.2698	17.6514
50	0.09409	55.10938	0.272	19.982

Table 4
Estimates of fatalities expected to be exceeded once every T years.

5 years			10 years			25 years		
μ	Fat.MLE	Fat.MOM	μ	Fat.MLE	Fat.MOM	μ	Fat.MLE	Fat.MOM
40	133	132	40	177	177	40	236	238
45	140	139	45	184	183	45	243	245
50	141	141	50	186	185	50	247	248

traffic management and flying conditions permit to maintain the model used in this study.

4. Uncertainty quantification

In the area of insurers and reinsurers, the management and prediction of extreme risk events like aviation accidents fatalities is very important. However, most mathematical models do not provide a perfect representation of reality; the adequacy of model and inference of models parameters are often uncertain. Thus, taking into account assessing uncertainties is very important when we extrapolate the observed data.

In literature, several methods (graphical and numerical) have been proposed to assess different sources of uncertainties in the extreme value statistics, for more details see [Wehner \(2010\)](#), [Scarrot and Macdonald \(2012\)](#), [Zhang et al. \(2015\)](#) and [Wendy and Noriszura \(2012\)](#). In this work, we focus on uncertainties in model adequacy and model parameters due to threshold selection.

4.1. Uncertainty from model adequacy

We use quantile plot and probability plot to check the suitability of the fitted GPD to the set of extracted threshold exceedances. The analysis of the previous section has shown that the standard error of shape parameter in MOM and MLE are less than the standard error in PWM; For that reason, we consider MOM and MLE methods in our analysis. We use the graphical approach to quantify uncertainty from model adequacy; the QQ plots and PP plots are done for MLE and MOM estimates for the threshold range 40 50, specifically for the threshold values 40, 45 and 50.

[Figs. 5 and 6](#) show the QQ plot and PP plot for thresholds 40 and 45 in MLE (left) and MOM (right); in MLE, QQ plot has a clear deviation from 45 degree line. In MOM, QQ plot is 45 degree line and PP plot is 45 degree line in both methods.

[Fig. 7](#) shows that QQ plot and PP plot are 45 degree line in both methods. The synthesis of graphical analysis shows that all QQ plots and PP plots for MOM are 45 degree line; this means that MOM is more suitable in this work. In the following, all the estimates are done with MOM estimator.

4.2. Uncertainty from estimate of the model parameters

The main source of uncertainty in quantifying extreme events such extreme aviation fatalities with GPD is the choice of the threshold. In this work, we use graphical method and intensive parameter fitted with MLE, MOM and PWM to select the appropriate range for threshold and to obtain the optimal threshold. There are some tools to measure the uncertainty in fitted GPD like coefficient of variation (CV), standard error (SE), confidence level of the parameter of interest. In our approach, for each potential threshold, standard error of shape and scale parameter are computed. The results obtained suggest that MLE and MOM are suitable than PWM to fit GPD parameter. In the previous section, [Table 4](#) summarizes the estimates of an T return level for different levels and two selected methods. Here, in order to quantify the uncertainties in parameter estimates, we investigate the basis of their standard error, coefficient of variation and confidence intervals. However, a standard 95% confidence interval for level x_T depends on the asymptotic normality of \hat{x}_T . This asymptotic normality assumption is not always proven out in practice for data with small observations, see

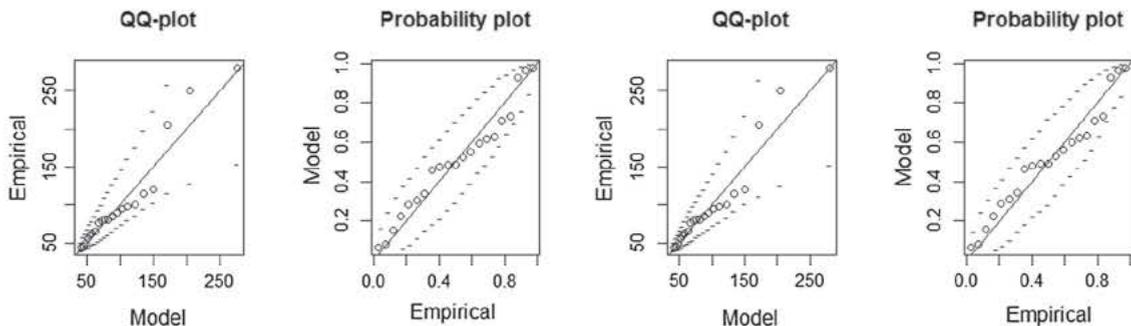


Fig. 5. QQ-plot and PP-plot for MLE (right) and MOM (left) with thres. = 40.

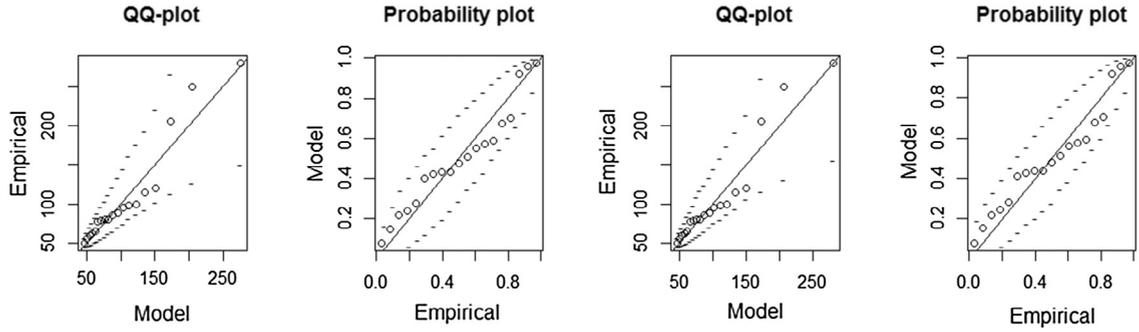


Fig. 6. QQ-plot and PP-plot for MLE (right) and MOM (left) with $\text{thres.} = 45$.

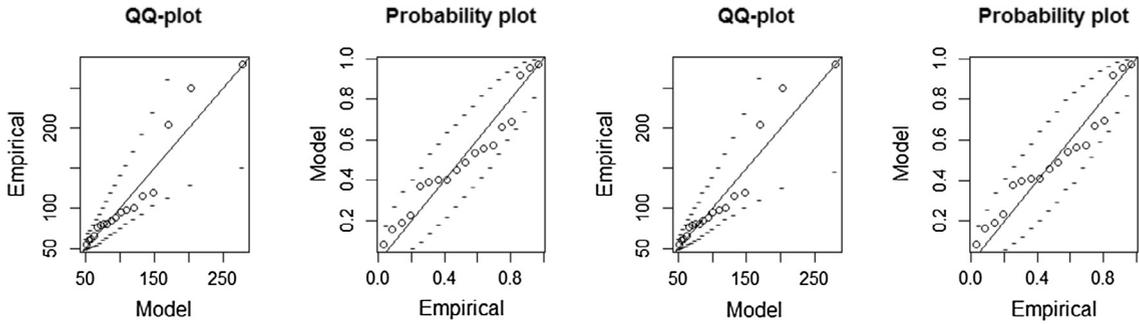


Fig. 7. QQ-plot and PP-plot for MLE (right) and MOM (left) with $\text{thres.} = 50$.

Carpenter and Bithel (2000) and Naess and Clausen (2001). We want to construct a confidence interval without taking into account this assumption. Hence, we use nonparametric resampling (nonparametric bootstrap) which makes no assumption concerning of model. Our data is assumed to be a vector z_{obs} of n independent observations, and we are interested in a confidence interval for $\theta_{z_{obs}}$. The general algorithm for a nonparametric bootstrap is as follows:

1. Sample n observations randomly with replacement from z_{obs} to obtain a bootstrap data set, denoted Z^* . However, in our case the original data set consists of number of observations with exceedances above the selected threshold. A random sample size number of exceedances is selected with replacement from the original data set to obtain a bootstrap data set. From this bootstrap data set MOM estimates of parameters are calculated and using these estimates the return level x_T is determined for each return period $T = 5, T = 10$ and $T = 25$.
2. Calculate the bootstrap version of the statistic of interest, $\theta^* = \theta(Z^*)$.
3. Repeat steps 1 and 2 a large number of times, say B , to obtain an estimate of the bootstrap distribution.

In our simulation, $B = 1000$, the statistic of interest is return level x_T , shape parameter ξ and scale parameter σ .

The procedure is repeated 1000 times; for each return period (T) we have 1000 values of return level. Using this set of 1000 values the mean return fatalities, standard error, coefficient of variation, minimum and maximum fatalities are calculated for thresholds 40, 45 and 50. The results are summarized in Tables 5 7.

The examination of results in Tables 5 7 shows that the standard error, the coefficient of variation, the minimum and the maximum fatalities for return level increase with the return period. This means that the uncertainties increase with the return period. We also note that there is no significant difference between the results of thresholds 45 and 50; this implies that the model gives almost similar results for any threshold value between 45 and 50. According to model adequacy and nonparametric bootstrap approach, we can say that the optimal threshold is around 45.

Results in Table 8 show that the root mean square error (rmse) for the parameter estimates (shape and scale) are almost the same.

Table 5

Mean fatalities, standard error, coefficient variation, min fatalities, max fatalities from bootstrap resampling (threshold 40).

Ret.period	Mean fat.	Threshold 40			
		S.E	C.V	Min fat.	Max fat.
5	131	21.32474	16.23269	81	207
10	174	29.47477	16.92119	92	273
25	227	44.64727	19.65789	100	343

Table 6

Mean fatalities, standard error, coefficient variation, minimum fatalities, maximum fatalities from bootstrap resampling (threshold 45).

Ret.period	Mean fat.	Threshold 45			
		S.E	C.V	Min fat	Max fat.
5	137	21.32474	16.23269	91	224
10	179	33.2863	18.65247	94	289
25	231	45.7403	19.76416	94	362

Table 7

Mean fatalities, standard error, coefficient variation, min fatalities, max fatalities from bootstrap resampling (threshold 50).

Ret.period	Mean fat.	Threshold 50			
		S.E	C.V	Min.fat.	Max.fata
5	137	22.11979	16.114	86	247
10	179	32.6103	18.25288	94	278
25	232	47.86963	20.63656	93	359

Table 8

rmse on parameters and return levels (5, 10 and 25 years).

GPD	$rmse(\sigma)$	$rmse(\xi)$	$rmse(x_5)$	$rmse(x_{10})$	$rmse(x_{25})$
$GPD(40,58.7971,0.0458)$	18.42124	0.2740146	20.60037	32.48511	46.36451
$GPD(45,60.26450,0.03779)$	20.11506	0.3053149	24.30641	33.95647	47.77649
$GPD(50,56.67434,0.06837)$	19.68647	0.2385574	26.24385	34.12842	49.21625

The same goes for the return level estimates. This analysis confirms the results in [Tables 5](#) [7](#); thus we can observe that there is no misspecification problem due to the threshold range selection of the model around threshold value 45.

4.3. Comparison between return level with original data and simulated data

In the resampling procedure, the original data set consists of number of exceedances above the thresholds 40, 45 and 50. A sample of size of exceedances number is selected with replacement from the original data set to obtain a bootstrap data set. The entire number of fatalities above the selected threshold changes with replacement. Thus, the expected values and the mean values from the bootstrap must be in 95% bootstrap confidence intervals to obtain accurate results.

According to [Tables 5](#), [6](#), [7](#) and [9](#), we observe that the mean fatalities and expected fatalities are close to 95% bootstrap confidence intervals for different return periods and thresholds.

5. Discussion and conclusion

The main source of uncertainty in estimation of GPD is due to the optimal threshold selection and method of estimation. Many authors have investigated in this sense, [Bermudez and Kotz \(2010\)](#) carried out an extensive study on several types of methods for estimating the GPD parameters; the authors argued that the success of GPD on a set data depends substantially on the process of parameter estimate, [Castillo and Daoudi \(2009\)](#) provide precise arguments to explain the anomalous behavior of the likelihood surface when sampling the GPD for small or moderate samples, [Far and Wahab \(2016\)](#) propose an approach to specify the most suitable threshold value for POT model which is called hybrid method, [Zhang et al. \(2015\)](#) focus on performing POT method from quantification of statistical uncertainties, etc. However, there is no universal approach to get the accurate threshold; in all existing methods there is a certain susceptibility.

The assessment of risks of extreme GA accident is very important because most of U.S. aviation accidents are in GA area. Hence,

Table 9Estimated fatalities expected to be exceeded once every T years based on GPD model for different potential threshold values with 95% bootstrap confidence intervals in brackets.

T years		$\mu = 40$	Threshold $\mu = 45$	$\mu = 50$
		5	132[97,178]	139[99,189]
10		177[111,239]	183[109,243]	185[111,250]
25		238[118,311]	245[123,322]	248[118,326]

the question of the prediction of future fatalities arising from GA are a real challenge; but the perfect representation of reality in U.S. GA accident fatalities is a hard work. We propose from extreme value statistics an approach based on Generalized Pareto Distribution, the prediction of the number of fatalities resulting in extreme GA accidents in the future operations. This approach could be applied to any retrospective set of data to other time periods; however, in this case we must determine the new threshold and other parameters of model 4. The fitted model permits to have approximately a probable number of fatalities from GA accidents which is expected to be exceeded once in a certain period in years.

In order to quantify the uncertainties in the estimates, we use the nonparametric bootstrap approach resampling; to check the adequacy of GPD model, we use the graphical approach. The results indicate that our nonparametric bootstrap approach provides accurate estimate of the uncertainty associated with parameter estimation. The probable number of fatalities from U.S. GA accident is given by the fitted model with the following assumptions: there is no change in the technology, tools or equipments used and the volume of air traffic. These assumptions are questionable in the aviation field according to the improvement of technology and aircraft operations. This can be regarded as another source of uncertainty in the prediction of fatalities; the future challenge of this work is to take into account this uncertainty using Prognostic and Health Management approach.

Acknowledgements

The authors wish to express gratitude to the Malian Government for administrative procedures. Particular thanks go to French Embassy in Bamako for the funding of research residency.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.tra.2018.01.022>.

References

- Airbus, 2016. Commercial Aviation Accidents, A Statistical analysis, 1958–2016. Airbus, Toulouse.
- Bermudez, Z., Kotz, S., 2010. Parameter estimation of the generalized Pareto distribution – Part II. *J. Statist. Plann. Inference* 140, 1374–1388.
- Boeing, 2016. Statistical Summary of Commercial Jet Airplane Accidents Worldwide Operations, 1959–2013. Aviation Safety-Boeing Commercial Airplanes, Washington.
- CAA, 2013. Global fatal accident review. CAA, Norwich.
- Carpenter, J., Bithel, J., 2000. Bootstrap confidence intervals when, which, what? a practical guide for medical statisticians. *Statist. Med.* 19, 1141–1164.
- Castillo, J., Daoudi, J., 2009. Estimation of generalized Pareto distribution. *Statist. Probab. Lett.* 79 (5), 684–693.
- Castillo, E., Hadi, A., Balakrishnan, N., Sarabia, M., 2005. *Extreme Value and Related Models with Applications in Engineering and the Sciences*. Wiley, Wiley-Interscience.
- Coles, S., 2001. *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag, New York.
- Crane, D., 1997. *Dictionary of Aeronautical terms*. Aviation Supplies-Academics, Baltimore.
- Das, A., Das, K., 2012. The generalized Pareto distribution and threshold analysis of normalized hurricane damage in the united states gulf coast. *Statist. Meet. (JSM)*.
- Dey, A., Das, K., 2016. Modeling extreme hurricane using the generalized Pareto distribution. *Amer. J. Math. Manage. Sci.* 35 (1), 55–66.
- Dupuis, D., Tsao, M., 1998. A hybrid estimator for gpd and extreme value distributions. *Commun. Stat.-Theory Methods* 27 (4), 925–941.
- Edwards, A., Das, K., 2016. Using the statistical approach to model natural disasters. *Am. J. Undergrad. Res.* 13 (2), 87–104.
- Farah, H., Azevedo, C., 1952. Safety analysis of passing maneuvers using extreme value theory. *IATSS Res.* 41, 12–21.
- Far, S., Wahab, A., 2016. Evaluation of peaks over threshold method. *Ocean Sci.*
- Fultz, A., Ashley, W., 2016. Fatal weather-related general aviation accidents in the united states. *Phys. Geogr.*
- Hosking, J., Wallis, J., 1987. Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics* 29 (3), 339–349.
- Knetch, W., 2015. Civil Aerospace Medical Institute. FAA, Washington.
- Mackay, E., Challenor, P., Bahaj, A., 2014. A comparison of estimators for the generalized Pareto distribution. *Ocean Eng.* 38, 1338–1346.
- Naess, A., Clausen, P., 2001. Combination of the peaks-over-threshold and bootstrapping methods for extreme value prediction. *Struct. Saf.* 23, 315–330.
- NTS, 2016. National Transportation Statistics. U.S. Department of Transportation, Washington.
- NTSB, 2011. Review of U.S. Civil aviation accidents calendar year. National Transportation Safety Board, Washington.
- Pickands, J., 1975. Statistical inference using extreme order statistics. *Ann. Stat.* 3 (1), 119–131.
- Ribatet, M., 2012. Pot: Generalized Pareto distribution and peaks over threshold, r package version 1.1-3. CRAN.
- Romeijnnders, W., Teunter, R., Jaarveld, W., 2012. A two step method for forecasting spare parts demand using information on component repairs. *Eur. J. Oper. Res.* 220 (2), 386–393.
- Scarrot, C., Macdonald, A., 2012. A review of extreme value threshold estimation and uncertainty quantification. *REVSTAT Statist. J.* 10 (1), 33–60.
- Schmidt, Z., Zhou, X., Toulemonde, F., 2014. A peaks over threshold of extreme traffic load effects on bridges. HAL.
- Sha, Z., Rommert, D., Willem, V., Rex, W., Alex, J., 2017. An improved method for forecasting spare parts demand using extreme value theory. *Eur. J. Oper. Res.* 261 (1), 161–181.
- Shetty, K., Hansman, R., 2012. *Current and Historical Trends in General Aviation in the United States*. MIT ICAT, Cambridge.
- Sobieralski, J., 2012. The cost of general aviation accidents in the U.S. transportation research Part a: Policy and practice. *REVSTAT Statist. J.* 47, 19–27.
- Wehner, W., 2010. Sources of uncertainty in the extreme value statistics of climate data. *Extremes* 19, 205–217.
- Wendy, S., Noriszura, I., 2012. Analysis of t-return level for partial duration rainfall series. *Sains Malaysiana* 41, 1389–1401.
- Yongcheng, Q., 2008. Bootstrap and empirical likelihood methods in extremes. *Extremes* 11, 81–97.
- Zhang, Y., Cao, Y., Dai, J., 2015. Quantification of statistical uncertainties in performing the peaks over threshold method. *J. Mar. Sci. Technol.* 23 (5), 717–726.