Aerosp. Sci. Technol. 142 (2023) 108646

Contents lists available at ScienceDirect

Aerospace Science and Technology

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

Data-driven mid-air collision risk modelling using extreme-value theory

Benoit Figuet [a,b,*], Raphael Monstein [a,b], Manuel Waltert [b], Jérôme Morio [c]

[a] Centre for Aviation, Zurich University of Applied Sciences, Technikumstrasse 71, Winterthur, 8400, Switzerland
[b] SkAI Data Services, Querstrasse 17, Zurich, 8050, Switzerland
[c] ONERA/DTIS, Université de Toulouse, F-31055 Toulouse, France

ARTICLE INFO

Communicated by Christian Circi

Keywords:
Aviation safety
Collision risk modelling
Extreme-value theory
Peaks Over Threshold
Monte Carlo simulation

ABSTRACT

Mid-air collision risk estimation is crucial for maintaining aviation safety and improving the efficiency of air traffic procedures. This paper introduces a novel, data-driven methodology for estimating the probability of mid-air collisions between aircraft by combining Monte Carlo simulation and the Peaks Over Threshold approach from Extreme Value Theory. This innovative approach has substantial advantages over traditional methods. Firstly, it reduces the number of assumptions about the traffic flow compared to traditional analytical methods. In fact, data-driven techniques require fewer assumptions, as they inherently capture the structures of the traffic flow within the underlying data. Secondly, it converges faster than methods based on crude Monte Carlo simulation. Notably, by employing Extreme Value Theory, this approach enables efficient evaluation of low-probabilities, which are commonly found in collision risk modelling. The effectiveness of the proposed methodology is demonstrated through estimating the probability of a mid-air collision in a real-world practical example. The case study investigates the risk of collisions between departures and go-arounds in the terminal airspace at Zurich Airport, highlighting the potential for improved safety and efficiency in air traffic management.

1. Introduction

Civil Aviation Authorities (CAA) and Air Navigation Service Provides (ANSP) are responsible to operate the airspace in their charge safely and efficiently. This responsibility includes the management of risk of Mid-Air-Collisions (MAC), which can be estimated quantitatively. The estimated MAC probabilities of a certain operation in an airspace can, thus, be compared with an acceptable risk level, called the Target Level of Safety (TLS). The total risk of a fatal accident of an entire flight is approximately 1×10^{-7} [1]. This total risk budget is usually subdivided into the different flight phases and causes. For example, in en-route airspace, the typical TLS for MAC risks is set at around 1×10^{-8} fatal accidents per flight hour [2], with one collision accounting for two fatal accidents. To evaluate the safety of new separation standards in en-route airspace, the International Civil Aviation Organization (ICAO) further breaks down this risk into separate longitudinal, lateral, and vertical risk values, each of which corresponding to a probability of 5×10^{-9} [3]. For terminal airspace, which refers to airspace adjacent to aerodromes, the TLS for collisions is less well-defined in the literature. Moreover, the risk is not defined "per flight hour" but rather "per movement", i.e., per arrival or departure, depending on the situation. The TLS allocated to MAC in terminal airspace is commonly set at 1×10^{-8} fatal accidents per movement, which corresponds to a collision probability of 5×10^{-9} per aircraft. For a comprehensive overview on the TLS applied in aviation, the reader is further referred to the works of Lin et al. [4], Eurocontrol [2].

To estimate the risk of a MAC, multiple analytical collision risk models have been developed and established over the years. Amongst the most influential are the analytical models of Reich [5] and of Hsu [6]. These models (i) assume traffic distributions along predefined routes based on required or observed navigational performance [7]; (ii) consider distributions for normal (core) and abnormal (tail) operations, e.g., to define separation standards for parallel tracks, or to model speeds and flown headings of aircraft [8]; (iii) take into account operational errors, such as large deviations at given angles [9]; and (iv) consider traffic densities in discrete grid cells. Variants of these models have historically been applied by the ICAO and by CAAs to define new separation standards. Mitici and Blom [10] provides a general overview of the different mathematical collision risk models.

These analytical models are effective for en-route operations with well-structured traffic flows; however, analytical models are often too simplistic to address complex traffic patterns typically found in termi-

* Corresponding author at: Centre for Aviation, Zurich University of Applied Sciences, Technikumstrasse 71, Winterthur, 8400, Switzerland.
E-mail address: benoit.figuet@zhaw.ch (B. Figuet).

https://doi.org/10.1016/j.ast.2023.108646
Received 26 June 2023; Received in revised form 4 September 2023; Accepted 20 September 2023
Available online 25 September 2023
1270-9638/© 2023 The Author(s). Published by Elsevier Masson SAS. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

nal manoeuvring areas (TMAs) and control zones (CTRs) adjacent to aerodromes, collectively referred to as terminal airspace (TA). Indeed, these models have not been designed to model flights that are turning, climbing, or descending. To circumvent this limitation, Figuet [11] proposed an approach where flight paths are modelled by piecewise linear segments.

The availability of large amounts of data, combined with advancements in computational power to process this data, has facilitated the development of data-driven approaches to mid-air collision risk modelling [12–15]. Such data-driven approaches are capable of better addressing the limitations encountered in analytical methods, as the underlying distributions are inherently present in the data. Unfortunately, existing data-driven methods still heavily rely on Monte Carlo simulations, which can be computationally demanding and necessitate a large number of simulations to accurately estimate the low probabilities associated with collision risks, in some circumstances up to the point of becoming infeasible.

To address limitations in existing data-driven collision risk models, we propose a novel approach that combines Monte Carlo simulations with Extreme Value Theory (EVT) to assess the probability of mid-air collision. EVT provides a well-established framework for modelling the tail of a probability distribution, making it particularly well-suited for modelling rare events, such as mid-air collisions. Indeed, EVT has been successfully employed in other fields, such as finance [16], insurance [17] and civil engineering, where EVT is used to model rare events such as extreme market fluctuations, or natural events such as extreme wave heights [18]. In terms of aviation-related applications, [7] shows how EVT can be used to model large aircraft trajectory deviations.

This paper presents a novel data-driven methodology to estimate the MAC probabilities by combining Monte Carlo simulation with EVT methodologies. This innovative methodology offers several key advantages over traditional data-driven collision risk modelling methods. Firstly, it inherently captures assumptions about traffic flows within the data, eliminating the need for introducing additional assumptions about traffic patterns. Secondly, the incorporation of EVT allows for a more efficient convergence in estimating low-probability events, making the methodology particularly effective for assessing the probability of rare occurrences such as mid-air collisions. Overall, the proposed approach aims to provide a robust and practical solution for mid-air collision risk modelling in a wide range of contexts in aviation and other disciplines.

The structure of this paper is organized as follows: After this introduction, Section 2 provides background information on EVT. The proposed method, which consists of a combination of Monte Carlo simulations and EVT, is presented in Section 3. Additionally, an approach to quantify the uncertainty of the method is shown. Subsequently, the model is applied to a real-world example in Section 4. To this end, the collision probability between an eastbound take-off (TO) on runway 16 and a simultaneous go-around (GA) on runway 14 at Zurich Airport in Switzerland is estimated. Section 5 discusses both the proposed method as well as the practical example in terms of their strength, limitations, and potential for enhancing aviation safety. Finally, Section 6 concludes this paper by summarizing the key findings and identifying both the implications of this work on the field of MAC risk modelling and potential for future research.

2. Background information

This section provides background information on (i) Extreme Value Theory and (ii) the concept of Peaks Over Threshold. This information provides the basis for the data-driven collision risk modelling methodology presented in Section 3.

Extreme value theory Irrespective of the field, risk management is about understanding extremes. Although the Central Limit Theorem (CLT) is commonly employed to model averages, it primarily addresses the behaviour of sums and averages of random variables as their quantity increases, ultimately converging to a normal distribution. However, when the interest lies in the tail behaviour of distributions or in the study of extreme events, the CLT may fall short [19].

Extreme Value Theory (EVT) was explicitly developed to model the probability of extreme events by finding reliable estimates of their frequency. Thus, EVT is concerned with the distribution of the partial maxima $M_n = \max(X_1, \dots, X_n)$, where $X_1, \dots, X_n$ is a sequence of independent and identically distributed (i.i.d.) random variables with common Cumulative Distribution Function (CDF) F. In other words, EVT aims to describe a sequence of maxima, or rare events. In this case, the CDF of the sample maximum, M_n, is given by $\mathrm{P}(M_n \leq x) = \mathrm{P}(X_1 \leq x, \dots, X_n \leq x)$. Most of the time, the distribution function F is unknown, however, the *Fisher-Tippett-Gnedenko* theorem [20,21]) provides an asymptotic result: Suppose there exist two sequences of real numbers $a_n > 0$ and $b_n \in \mathbb{R}$ such that the following limit converges to a non-degenerate distribution function G:

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{\max\{X_1, \dots, X_n\} - b_n}{a_n} \leq x\right) = G(x), \tag{1}$$

where the limit distribution G belongs to the Generalized Extreme Value (GEV) distribution family, more precisely to either the Gumbel, the Fréchet, or the Weibull family. In practical terms, if the normalized maxima, i.e., location-scale transformed maxima, of i.i.d. random variables converges to a non-degenerate distribution G, then the limiting distribution must be a location-scale transformed GEV distribution.

Peaks Over Threshold The Peaks Over Threshold (POT) method is used to statistically describe the samples that exceed a user defined threshold u, i.e., the samples that are above a threshold u that is seen as an extreme value, as illustrated on simulated data in Fig. 1, plots a) and b). The *Pickands–Balkema–De Haan* theorem, see Balkema and De Haan [22], Pickands III [23], states that given a sequence of i.i.d. random variables $(X_1, \dots, X_n)$, their conditional excess distribution $F_u(y) = \mathrm{P}(X - u \leq y \mid X > u)$, is well approximated by the Generalized Pareto Distribution (GPD) $H_{\xi,\beta(u)}$

$$\mathrm{P}(X - u \leq y \mid X > u) = H_{\xi,\beta(u)}(y) = \begin{cases} 1 - \left(1 + \frac{\xi y}{\beta(u)}\right)^{-1/\xi} & \text{for } \xi \neq 0, \\ 1 - e^{-y/\beta(u)} & \text{for } \xi = 0, \end{cases} \tag{2}$$

with the scale parameter $\beta(u) > 0$. The support of $H_{\xi,\beta(u)}$ is $y > 0$ if the shape parameter ξ is $\xi \geq 0$, and $[0, \beta(u)/\xi]$ otherwise. The fit of such a distribution to the exceedances is illustrated in Fig. 1, plot c).

In the context of this paper, which aims to model the minimum distance between aircraft, the case where $\xi < 0$ is of particular significance. If $\xi < 0$, the GPD is bounded, as is the case when modelling the distance between two aircraft (the lower bound is 0, meaning the aircraft are at the same position). In this case, the GPD is referred to as a short-tailed Pareto Type II distribution and estimating its parameters is to find the values of ξ and β that maximize the log-likelihood under the constraint $u = \beta/\xi$.

Selecting an appropriate threshold value u is crucial when applying the POT method. If the chosen threshold value is too high, it results in fewer exceedances, leading to increased variance. On the other hand, selecting a threshold value that is too low introduces a bias in the estimation. Various techniques can be employed to determine an optimal threshold value. Among the most common methods are graphical approaches such as the Mean Excess Plot (MEP). As such, MEP offers a visual representation of the mean excess function $e_n(u)$ with respect to the threshold levels u. The empirical mean excess function is defined as

$$e_n(u) = \frac{\sum_{i=1}^{n} (X_i - u) I_{\{X_i > u\}}}{\sum_{i=1}^{n} I_{\{X_i > u\}}}, \quad u \leq X_{(n)}, \tag{3}$$

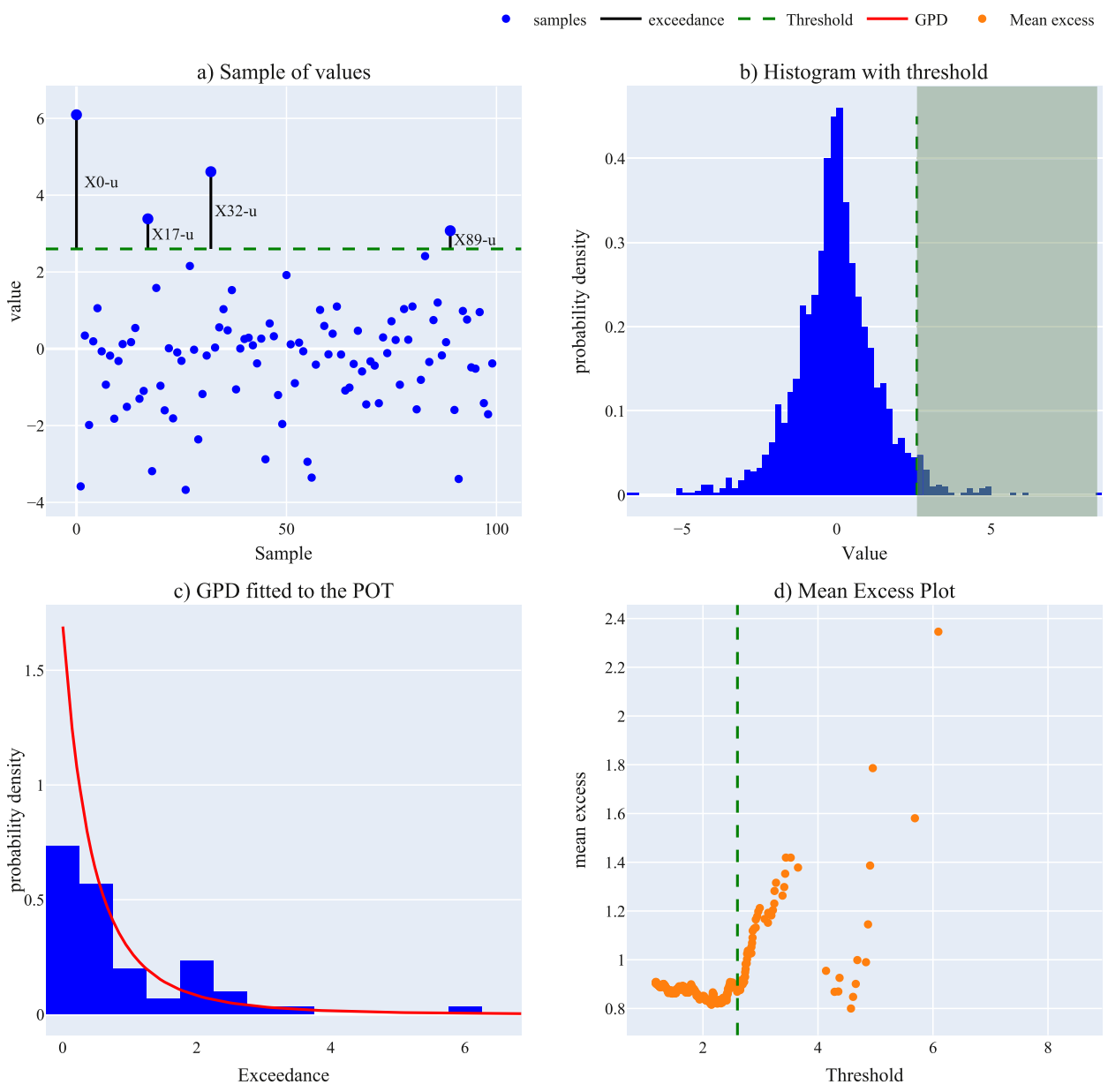


Fig. 1. Example of the Peaks Over Threshold method on simulated data sampled from a Gaussian-Laplace mixture distribution for illustration purposes. a) Sample of 100 values with the black lines indicating samples that exceed the threshold value u (green dashed line); b) Normalized histogram with the threshold indicated by the dashed green line and the area above the threshold shaded in green; c) Histogram of the exceedances over the threshold and the fitted generalized Pareto distribution (red line); d) Mean excess in function of threshold values. (For interpretation of the colours in the figure(s), the reader is referred to the web version of this article.)

with $X_1, ..., X_n$ being the i.i.d. sample and I the indicator function, i.e., $I_{\{X_i>u\}} = 1$ if $X_i > u$, otherwise $I_{\{X_i>u\}} = 0$. Subsequently, the actual MEP consists of the plot containing the points $\{(X_{(i)}, e_n(X_{(i)}) : 1 \leq i \leq n-1\}$, where $X_{(i)}$ denotes the i^{th} order statistic, i.e., the i^{th} smallest sample, and n the number of samples. A suitable value for the threshold u is usually taken as the smallest value where the MEP transitions into a roughly linear function. This linearity indicates that the GPD is a suitable approximation for the excess distribution above the threshold. An example for a MEP is shown in Fig. 1, plot d).

Besides MEP, Thompson et al. [24] propose an automated threshold selection based on the distribution of the difference of parameter estimates. For a comprehensive overview on the topic of threshold value selection, the reader is referred to Scarrott and MacDonald [25]. In practice, the choice of the threshold u is a difficult exercise and Embrechts et al. [16] recommends to conduct the analysis for multiple threshold values.

3. Methodology

In most collision risk models, aircraft are represented either as a cuboid or a cylinder. For the sake of simplicity, aircraft are represented in this paper as spheres with diameter λ, which refers to the wingspan of the aircraft, as illustrated in Fig. 2. Using a sphere to represent the aircraft is generally a more conservative approach due to the larger volume encompassed compared to a cylinder or a cuboid. If available, the appropriate diameter of the sphere can be estimated from historical data by extracting the average wingspan of the aircraft involved.

Having defined the geometrical shape to represent aircraft, the first step in modelling the collision probabilities is to simulate pairs of aircraft trajectories with Monte Carlo simulation runs. The pairs of trajectories can be obtained in various ways, including (i) from an air traffic simulator such as *Bluesky* [26], which relies on a performance model like *OpenAp* [27] or EUROCONTROLs *base of aircraft data (BADA)* [28],

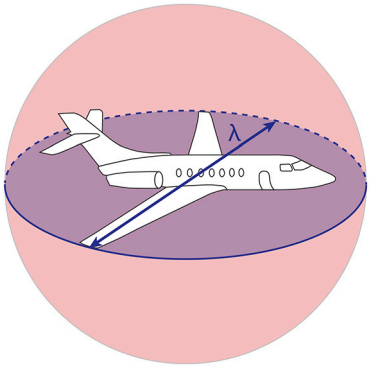


Fig. 2. Aircraft are represented by a sphere of diameter λ.

(ii) from a data-driven model as proposed by Krauth et al. [29], or (iii) by re-enacting observed aircraft trajectories, such as historical radar tracks. For each Monte Carlo simulation run, two aircraft trajectories are randomly selected. These two trajectories are replayed over time and the minimum observed three-dimensional Euclidian distance for the pair is extracted. This minimum distance is referred to as the spacing at the Closest Point of Approach (CPA), denoted d_{CPA}. By simulating numerous runs, a distribution of the spacings at the CPA is obtained.

In a second step, the POT approach is applied to the obtained distribution of the spacings at the CPA. In contrast to classical EVT, which focuses on modelling the behaviour of maximum values, our intention is to apply EVT to model the behaviour of minimum values, namely the minimum distance between the two aircraft obtained in each Monte Carlo run. The motivation behind this approach is to determine the probability of the spacing at CPA becoming smaller than an aircraft's diameter, i.e., the two aircraft collide. To achieve this, the following identity is used:

$$\min(X_1, \dots, X_n) = -\max(-X_1, \dots, -X_n). \tag{4}$$

As a result, the d_{CPA} values need to be converted into negative values and a collision is defined as $-d_{\text{CPA}} \geq -\lambda$.

To select an appropriate value for threshold u, we employ the MEP method. As mentioned in Section 2, we plot the mean excess function for different threshold values and select the threshold at the minimum value where the curve begins to exhibit linear behaviour.

Once a threshold value u is determined, we extract the CPA spacing exceedances over the threshold u for which we fit a GPD. Then, the parameters of the GPD are estimated using the Maximum Likelihood Estimation (MLE) technique. We assume that a perfect overlap ($d_{\text{CPA}} = 0$) is possible, as this methodology is intended to be used for scenarios where collisions can occur. This implies that the support of the GPD is bounded between u and 0. Consequently, we have the relationship $\beta = -\xi u$ with $\xi < 0$, and hence, only ξ needs to be estimated. With the GPD parameters in place, the probability of collision can be computed. Using Equation (2) yields the probability of a collision, given that the spacing at CPA is lower than threshold value u:

$$\begin{aligned} \mathrm{P}(\mathrm{MAC} \mid -d_{\text{CPA}} > u) &= 1 - H_{\xi,\beta(u)}(u-\lambda) = \left(1 + \frac{\xi(u-\lambda)}{\beta(u)}\right)^{-1/\xi} \\ &= \left(1 + \frac{\xi(u-\lambda)}{-\xi u}\right)^{-1/\xi} = \left(\frac{\lambda}{u}\right)^{-1/\xi} \end{aligned} \tag{5}$$

Applying the chain rule to Equation (5) yields the probability of a collision:

$$\mathrm{P}(\mathrm{MAC}) = \mathrm{P}(\mathrm{MAC} \mid -d_{\text{CPA}} > u) \cdot \mathrm{P}(-d_{\text{CPA}} > u) \approx \left(\frac{\lambda}{u}\right)^{-1/\xi} \cdot \frac{N}{n} \tag{6}$$

with n the number of simulated pairs, N the number of exceedances for the chosen threshold value u, and λ the diameter of the sphere representing the aircraft.

In a third and final step, we estimate the uncertainty associated with the MAC probability calculations. Equation (6) is the product of two probabilities which are both estimated with a certain degree of uncertainty. The estimate $\hat{\mathrm{P}}(-d_{\text{CPA}} > u)$ of $\mathrm{P}(-d_{\text{CPA}} > u)$ is obtained from Bernoulli trials at which success occurs at $-d_{\text{CPA}} > u$. Consequently, the estimate of the probability of success can be described thanks to the CLT with a random variable following a normal distribution, where:

$$\hat{\mathrm{P}}(-d_{\text{CPA}} > u) \sim \mathcal{N}\left(\frac{N}{n}, \frac{N(n-N)}{n^3}\right). \tag{7}$$

The second source of uncertainty is due to the estimation of the GPD parameters, in our case solely the parameter ξ. The standard errors associated with the MLE are obtained from the square roots of the diagonal elements of the inverse of the observed Fisher Information matrix $\mathbf{I}(\hat{\xi}_{\text{MLE}})$, i.e., the Fisher Information matrix evaluated at the maximum likelihood estimates [30]. The Fisher Information $\mathbf{I}(\xi)$ is obtained by evaluating the Hessian of the log likelihood. In our case, ξ is a scalar and thus the variance corresponds to the inverse of the observed Fisher Information. The Maximum Likelihood Estimates being normally distributed, we get the following equation

$$\hat{\xi}_{\text{MLE}} \sim \mathcal{N}\left(\hat{\xi}_{\text{MLE}}, [\mathbf{I}(\hat{\xi}_{\text{MLE}})]^{-1}\right) \tag{8}$$

with $[\mathbf{I}(\hat{\xi}_{\text{MLE}})]^{-1}$ is the variance of the estimated shape parameter.

The distribution of the MAC probability estimate $\hat{\mathrm{P}}(\mathrm{MAC})$ can finally be estimated with Monte Carlo simulation by sampling random variables $\hat{\mathrm{P}}(-d_{\text{CPA}} > u)$ and $\hat{\xi}_{\text{MLE}}$ for a sufficient number of times:

$$\hat{\mathrm{P}}(\mathrm{MAC}) = \left(\frac{\lambda}{u}\right)^{-1/\hat{\xi}_{\text{MLE}}} \cdot \hat{\mathrm{P}}(-d_{\text{CPA}} > u). \tag{9}$$

The obtained distribution can subsequently be used to extract descriptive statistics on $\hat{\mathrm{P}}(\mathrm{MAC})$, such as the quantiles for different confidence intervals (CI).

4. Practical example

The following practical example shows how the methodology proposed in this paper can be used to compute the probability of mid-air collision in a real-world scenario. In this particular case, we quantify the probability of a collision between a departing aircraft and an aircraft simultaneously flying a go-around procedure at the airport of Zurich, Switzerland.

4.1. Scenario description

Various operating concepts are used at Zurich Airport. During the day, aircraft primarily take off from runway 28 and land on runway 14. Additionally, runway 16 is also used for both take-offs of heavier, long-haul aircraft. However, since the flight paths of go-arounds on runway 14 intersect with those of departing aircraft on runway 16, see Fig. 3, these runways cannot be operated independently. To address this issue, the following mitigation measure is currently applied by air traffic control: No departures on runway 16 are released if an aircraft approaching runway 14 is closer to the threshold than a certain cut-off distance. As illustrated in Fig. 4, the cut-off distance applied at Zurich Airport is 7 Nautical Miles (NM) (or 8NM if the departing aircraft is of type Airbus A380).

4.2. Assumptions

We make the assumption that every time an aircraft is cleared for take-off on runway 16, another aircraft is simultaneously approaching runway 14, with the approaching aircraft consistently positioned exactly 5NM from the runway threshold. The choice of 5NM instead of 7NM is motivated by the higher probability of collision at this distance, which, for the sake of this study, simplifies the estimation of the collision probability. Essentially, a shorter cut-off distance increases the

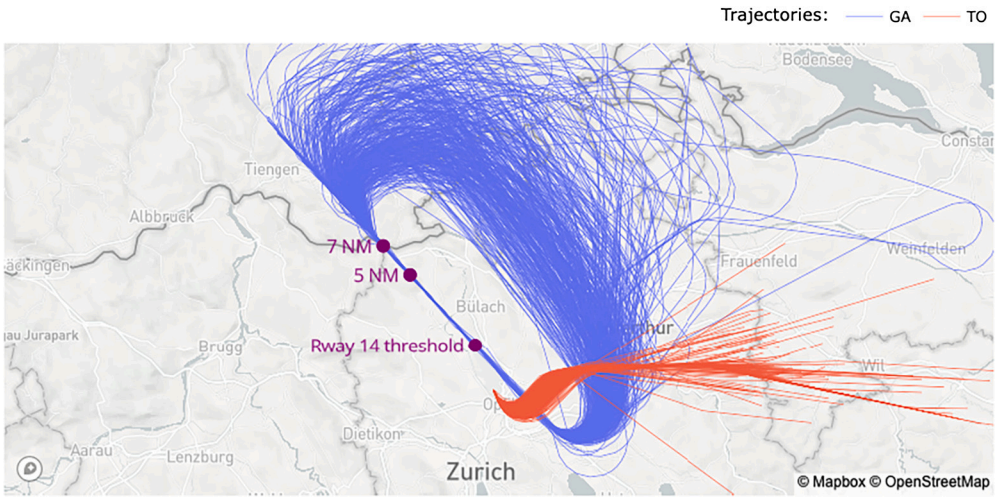
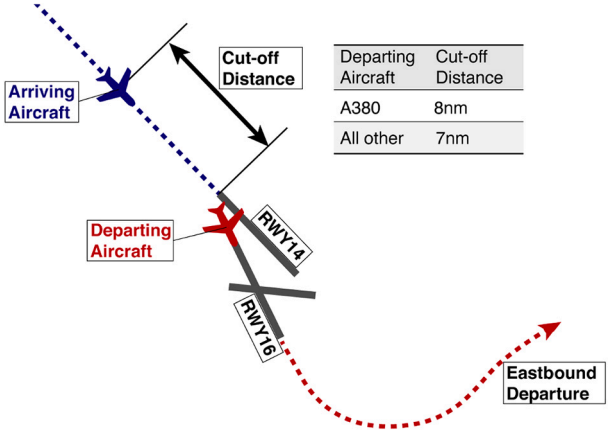


Fig. 3. Observed go-arounds on runway 14 (blue) and a sample of 400 eastbound departures from runway 16 (red) at Zurich Airport extracted from ADS-B data. The runway 14 threshold, along with markers at 5NM and 7NM from it, are displayed in purple.



Departing Aircraft	Cut-off Distance
A380	8nm
All other	7nm

Fig. 4. Current mitigation measures applied at Zurich Airport: No eastbound departures are released if an aircraft is approaching runway 14 is closer than 7NM from the threshold. If the departing aircraft is an A380, The cut-off distance is equal to 8NM.

risk, as it becomes more likely for the approaching aircraft to catch up with the departing one.

4.3. Data

This study uses Automatic Dependent Surveillance-Broadcast (ADS-B) data retrieved from the *OpenSky Network* [31], which contains, among other parameters, time-stamped position and altitude observations of aircraft. Go-arounds that have occurred since the year 2017 were retrieved and a total of 2,522 eastbound departures from the year 2022 were collected for analysis using the *traffic* library [32]. Out of approximately 800 GA collected, only 386 were retained for this study. Many of the retrieved GA did not adhere to the published procedure due to early left turns, which was the result of air traffic control interference. As the primary objective of this study is to assess the safety of the published procedure, only GA that followed the published procedure were considered for further analysis. Fig. 3 shows the retained GA on runway 14 and a random sample of 400 eastbound departures on runway 16.

As mentioned in Section 3, aircraft are represented by a sphere of diameter λ. To determine the value of λ, a frequency analysis of the aircraft types approaching runway 14 and departing from runway 16 as well as their wingspan was conducted. Fig. 5 summarizes the eight most commonly observed aircraft types and their associated wingspan. On average, a departing aircraft has a wingspan of 55 m, while an approaching aircraft, i.e. one flying a GA, has an average wingspan of 34 m. To maintain a conservative approach, it was decided to represent all aircraft with a sphere of diameter 55 m.

4.4. Monte Carlo simulation

The Monte Carlo method consists of simulating pairs of one GA and one departure. In this study, this is achieved not through traditional air traffic simulation, but rather by re-enacting previously observed and recorded ADS-B trajectories. After identifying both relevant departures and GAs, the system is simulated by sampling pairs of these observed trajectories. Combining each of the 386 GA with each of the 2,522 departures results in 973'492 combinations. A point of consideration is that the trajectories within a pair might have been observed during different atmospheric conditions, which may consequently impact the outcome. The simulation is conducted by initiating the GA aircraft at a distance of 5NM from the threshold of runway 14 the very moment the take-off roll of the aircraft departing from runway 16 is initiated. This enables us to mimic a conservative scenario where a departure is initiated while a landing aircraft is positioned exactly at the minimum cut-off distance allowed. In practice, the trajectory data quality of aircraft still being on the runway is insufficient for initiating the take-off at the threshold of runway 16. As an alternative, take-offs are initiated at the end of runway 16. For this reason, the time of take-off needs to be adjusted by one minute, which corresponds to roughly the average time it takes for an aircraft to reach the end of the runway after receiving a take-off clearance.

For each simulated aircraft pair, the three dimensional Euclidian distance at the CPA is retrieved and saved for later analysis using the POT method introduced in Section 2. For the distance at CPA to be accurate, all the trajectories are re-sampled at a resolution of 0.02 s. For illustrative purposes, Fig. 6 show four different simulation runs, where an aircraft pair consisting of a GA and a departure are replayed. Moreover, Fig. 7 depicts the observed distribution of the negative minimum spacings at CPA for the 973,492 GA/departure pairs. Since the aircraft is modelled as spheres of diameter $\lambda = 55$ m, a collision is described by a negative distance at CPA larger than -55 m, which is indicated with a red dashed line.

4.5. Peaks Over Threshold method

Following the CPA calculation, we delve into the extreme value analysis of the spacing at CPA to model. To this end, we analyze the behaviour at the tail of the CPA spacing distribution. The right plot

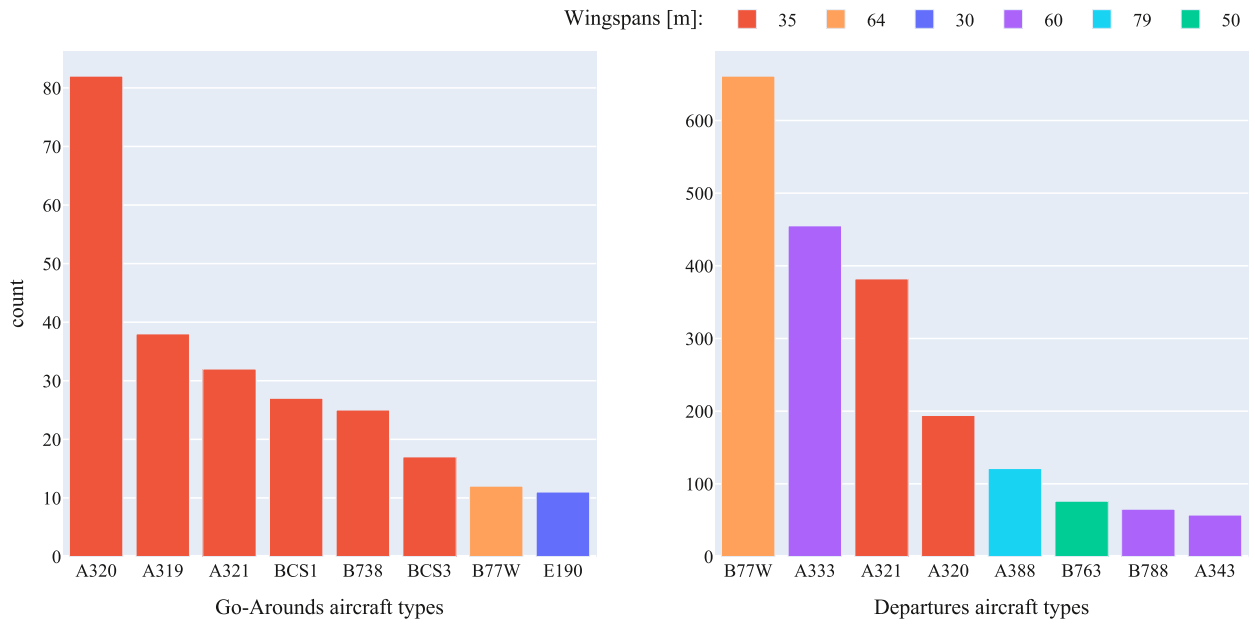


Fig. 5. The eight most commonly observed aircraft types flying go-arounds (left) and departures (right), and their associated wingspan. On runway 14, most of the GA are flown by narrow-body aircraft having an average wingspan of 34 m. Runway 16 is mostly used by wide-body aircraft with an average wingspan of 55 m.

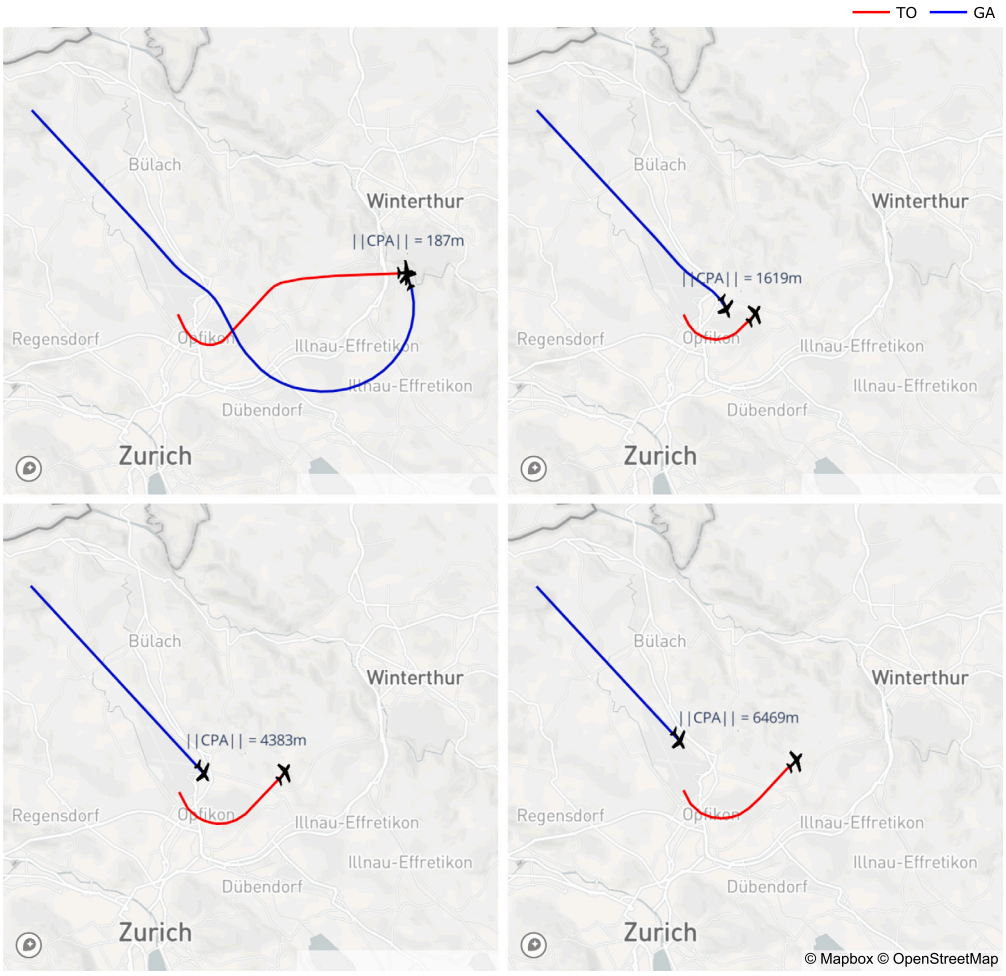


Fig. 6. Four pairs of replayed GA and departures illustrating various CPA spacings. The aircraft symbols illustrate the position of the respective aircraft at the CPA. Top left has a three-dimensional CPA spacing of 187 m; top right, 1,619 m; bottom left, 4,383 m; and bottom right, 6,469 m.

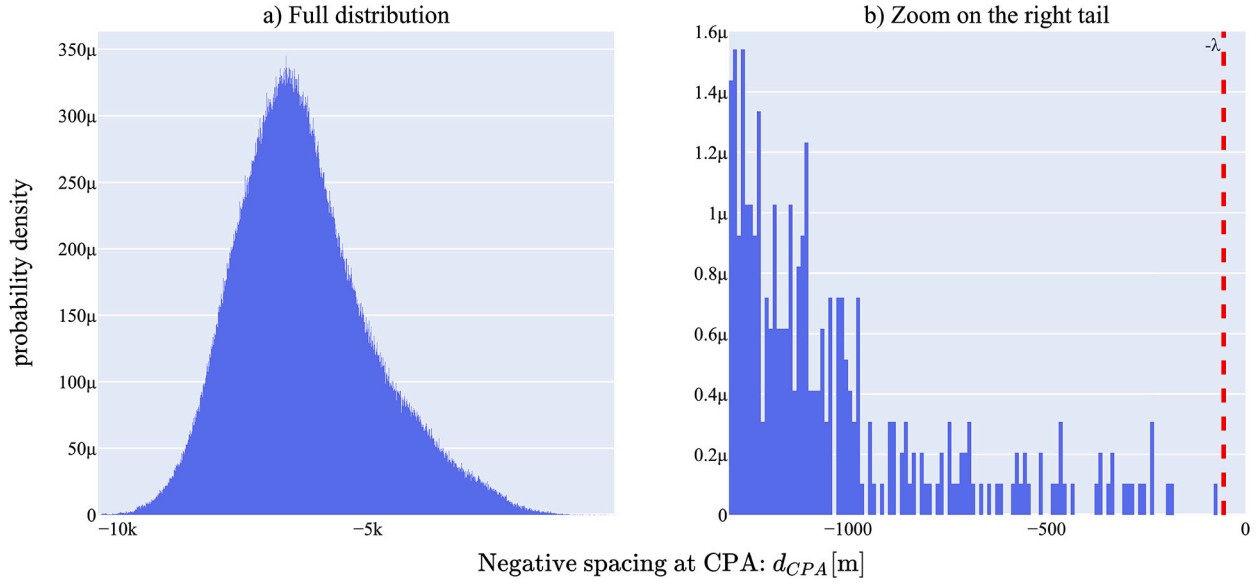


Fig. 7. The left plot shows the histogram of the negative spacings at CPA in meters for the 948,272 combinations of GA and departure pairs. The right plot is a zoom on the right tail of the distribution. The red line indicates the aircraft diameter λ. No collision was observed and the smallest observed negative spacing at CPA is -78 m.

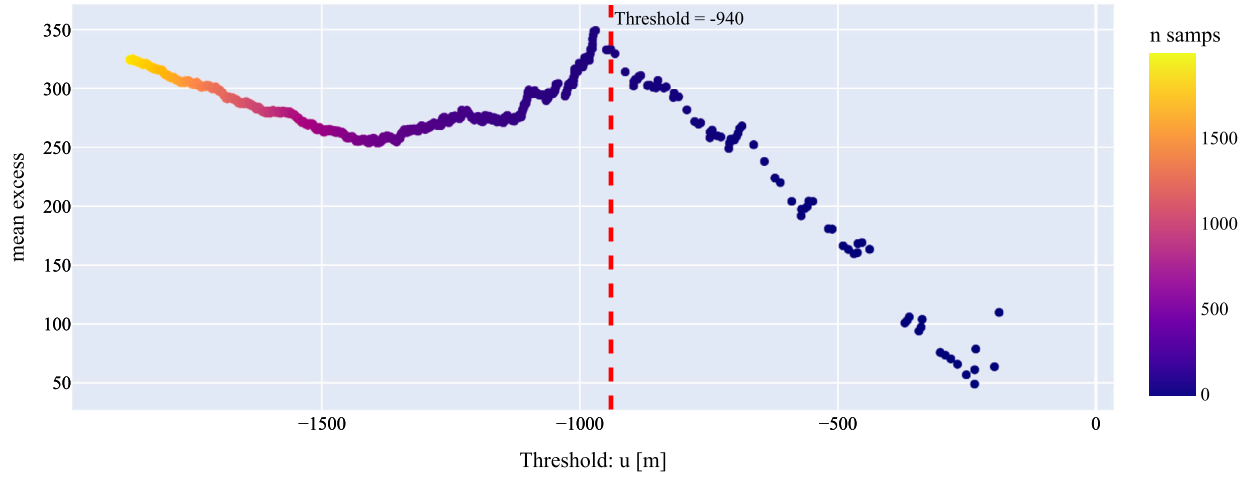


Fig. 8. Mean excess plot. The red line indicates the chosen threshold value u of -940 m.

in Fig. 7 provides a close-up view of this tail. Notably, no collisions were observed. Indeed, the closest spacing at CPA observed is 78 m. To conduct the EVT analysis, we employ the POT method, which requires selecting an appropriate value for threshold u. Fig. 8 shows the sample mean excess plot of the spacings at CPA. The threshold is chosen as the smallest value where the curve starts to exert a linear behaviour, i.e., at a value of −940 m, which results in 73 samples that are over the chosen threshold. As mentioned in Section 2, the choice of an appropriate threshold value can be difficult and it is recommended to repeat the analysis for multiple threshold values. This has been done for this study and the results are presented later in Section 4.8.

With the determined value for the threshold u, we proceed to fit a GPD to the tail (or more formally, to the CPA spacing exceedances over the threshold u). The value of the shape parameter of the GPD, ξ, was estimated using MLE to get $\hat{\xi}_{\text{MLE}}$, while the other parameter β is estimated by the relationship $\hat{\beta} = -\hat{\xi}_{\text{MLE}} \cdot u$. Fig. 9 depicts the histogram of the exceedances over the threshold u and the fitted GPD probability density function. Consequently, the fitted GPD is

$$H_{\xi,\beta(u)}(y) = \mathrm{P}(u - X \leq y \mid -d_{\text{CPA}} > u) \approx 1 - \left(1 + \hat{\xi}_{\text{MLE}} y / \hat{\beta}\right)^{-1/\hat{\xi}_{\text{MLE}}} \tag{10}$$

with $u = -940$, $\hat{\xi}_{\text{MLE}} = -0.534$, $\hat{\beta} = 501.93$, and d_{CPA} being the spacing at CPA.

4.6. Collision probabilities

With the GPD fitted to the tail, we employed Equation (6) to express the probability of a mid-air collision. However, in the practical example presented in this study, a collision can only occur if landing aircraft is performing a GA. Thus, the probability of a collision is conditioned on a GA occurring simultaneously:

$$\mathrm{P}(\text{MAC} \mid \text{GA}) \approx \frac{N}{n} \cdot \left(\frac{\lambda}{u}\right)^{-1/\hat{\xi}_{\text{MLE}}} \tag{11}$$

Using the chain rule, the probability of a collision is expressed as

$$\mathrm{P}(\text{MAC}) = \mathrm{P}(\text{MAC} \mid \text{GA}) \cdot \mathrm{P}(\text{GA}) \tag{12}$$

with P(GA) being the probability of a GA. Note that, as stated in Section 4.2, it is assumed that every time a take-off is released, an aircraft is on the approach of runway 14. P(GA) was estimated by using the dataset proposed by Monstein et al. [33], which contains a large number of both landings and GAs at Zurich Airport (amongst many others).

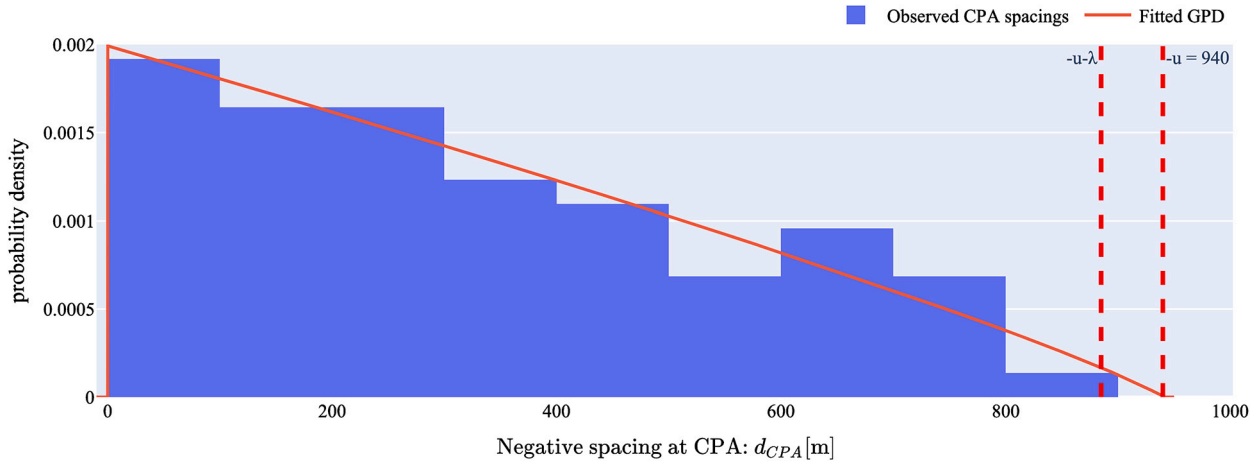


Fig. 9. Histogram of the CPA spacings exceedances over the threshold and its associated GPD fit. The collision probability given an exceedance can be obtained by integrating the GPD between the two dashed vertical red lines.

This dataset has been expanded to include landings from 2020 to 2023, allowing for a more precise estimation of the GA rate on runway 14 at Zurich Airport. Out of 181,933 recorded landings on runway 14, 598 resulted in a go-around. Consequently, the estimated GA probability is $\mathrm{P(GA)} \approx 598/181,933 \approx 3.3/1,000$.

Finally, we estimate the probability of a MAC as follows:

- Probability of a collision, given a simultaneous GA:

$$\begin{aligned}\mathrm{P(MAC \mid GA)} &\approx \frac{N}{n} \cdot \left(\frac{\lambda}{u}\right)^{-1/\hat{\xi}_{\mathrm{MLE}}} \\ &\approx \frac{73}{973,492} \left(\frac{\lambda}{u}\right)^{-1/\hat{\xi}_{\mathrm{MLE}}} \\ &\approx 7.5 \times 10^{-5} \cdot 4.91 \times 10^{-3} \\ &\approx 3.68 \times 10^{-7}\end{aligned}$$

- Probability of a GA:

$$\mathrm{P(GA)} \approx \frac{598}{181,933} \approx \frac{3.3}{1,000}$$

- Probability of a mid-air collision:

$$\mathrm{P(MAC)} = \mathrm{P(MAC \mid GA)} \cdot \mathrm{P(GA)} \approx 3.68 \times 10^{-7} \cdot \frac{3.3}{1,000} \approx 1.22 \times 10^{-9}$$

The estimated probability for a mid-air collision is $\hat{\mathrm{P}}(\mathrm{MAC}) = 1.22 \times 10^{-9}$. This implies that an eastbound departure from runway 16 has a probability of a mid-air collision of 1.22×10^{-9} if an approaching aircraft on runway 14 is situated at 5NM from the runway threshold when the take-off is released.

4.7. Uncertainty estimation

Compared to the uncertainty estimation presented in Section 3, the practical example considered has another source of uncertainty. Indeed, the additional uncertainty of the estimate for the probability of a GA, P(GA) must be accounted for. Since the estimate $\hat{\mathrm{P}}(\mathrm{GA})$ of P(GA) is obtained by Bernoulli trials, the same method can be used as for the uncertainty estimation of $\hat{\mathrm{P}}(-d_{\mathrm{CPA}} > u)$ in Equation (7):

$$\begin{aligned}\hat{\mathrm{P}}(\mathrm{GA}) &\sim \mathcal{N}\left(\frac{598}{181,933}, \frac{598(181,933 - 598)}{181,933^3}\right) \\ &\sim \mathcal{N}\left(3.3 \times 10^{-3}, 1.8 \times 10^{-8}\right)\end{aligned}$$

The probability estimate of having the CPA over the threshold u also follows a normal distribution, such that:

$$\begin{aligned}\hat{\mathrm{P}}(-d_{\mathrm{CPA}} > u) &\sim \mathcal{N}\left(\frac{N}{n}, \frac{N(n-N)}{n^3}\right) \\ &\sim \mathcal{N}\left(\frac{73}{978536}, \frac{73(978536 - 73)}{978536^3}\right) \\ &\sim \mathcal{N}\left(7.48 \times 10^{-5}, 7.62 \times 10^{-11}\right)\end{aligned}$$

The inverse Hessian of the log likelihood is equal to 3.91×10^{-3}, leading to the following expression for the uncertainty in ξ:

$$\hat{\xi}_{\mathrm{MLE}} \sim \mathcal{N}\left(-0.534, 3.91 \times 10^{-3}\right)$$

A Monte Carlo simulation with 100,000 runs was performed to obtain the distribution of probability estimate $\hat{\mathrm{P}}(\mathrm{MAC})$ and the histogram of the results is shown in Fig. 10. Fig. 10 also shows the values for the confidence intervals. It can be seen that for a 95% confidence interval, we have:

$$2.44 \times 10^{-10} \leq \mathrm{P(MAC)} \leq 3.35 \times 10^{-9}$$

It is worth highlighting that the confidence intervals obtained by EVT are smaller than the ones generated with other, simpler methods. For comparison, we analyzed the results of 978,536 Monte Carlo simulations, in which no collision was observed. The Rule of Three ([34]), states that for a sample of n subjects with zero occurrences of a certain event, a 95% confidence interval for the event's occurrence probability ranges from 0 to $3/n$. Applied to the case study, this means that the confidence interval for the probability of having a collision, given a GA, is

$$0 \leq \mathrm{P(MAC \mid GA)_{MC}} \leq 3.07 \times 10^{-6}.$$

The result of the Rule of Three is three times larger than the one obtained using the EVT for the same confidence level:

$$\mathrm{P(MAC \mid GA)_{EVT}} \leq 1.02 \times 10^{-6}$$

Comparing the proposed method with a classical Monte Carlo simulation-based approach shows that the proposed method is computationally more efficient. Using the Wilson score confidence interval [35], we can demonstrate that approximately 10.7 million Monte Carlo simulations are necessary to achieve a comparable confidence interval for the upper limit of the collision probability. Recalling that the method proposed in this study required fewer than one million Monte Carlo runs, makes the proposed method ten times more efficient. Given the relatively low number of observed GA, simulating 10.7 million Monte Carlo simulations would be very challenging with the data available.

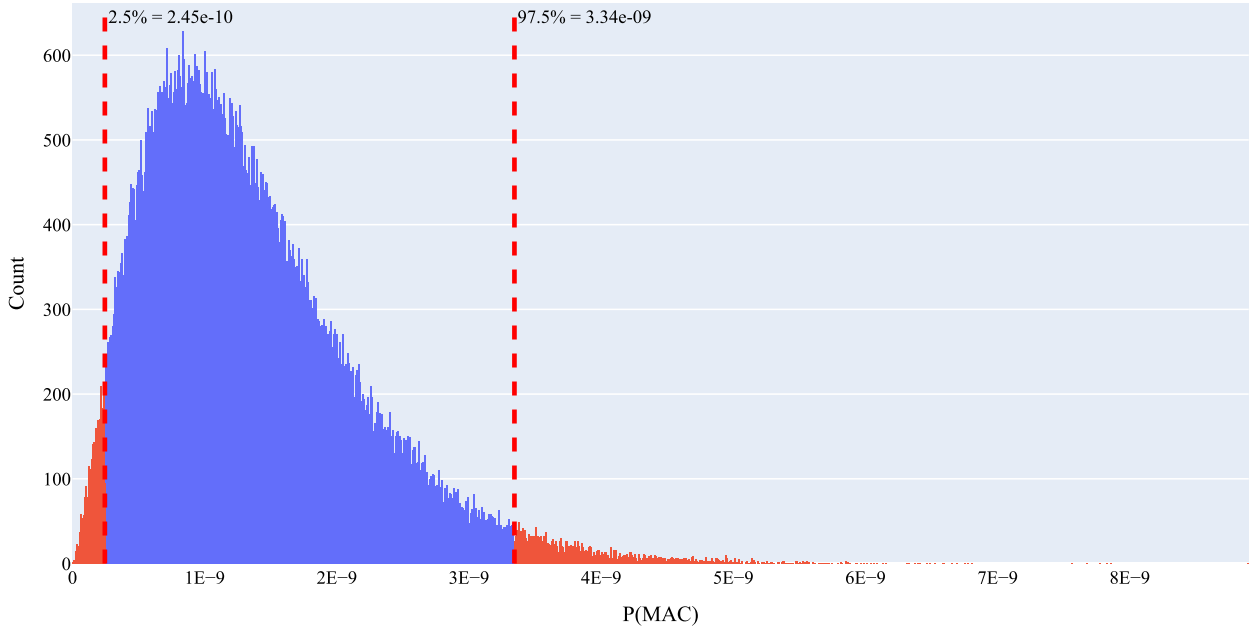


Fig. 10. Histogram of the $\hat{P}$(MAC) probabilities obtained by simulating 100,000 Monte Carlo runs. The lower and upper bounds corresponding to a 95% CI are indicated with red dashed lines.

4.8. Robustness of the estimates

As suggested by Embrechts et al. [16], the robustness of the collision probability estimates is assessed by using different values for the threshold u for the POT method. Fig. 11 illustrates the probability of collision for various threshold values, accompanied by the upper bound for a 95% confidence interval. Additionally, a Kolmogorov-Smirnov (KS) test has been performed to assess the goodness of fit for the GPD. The colour of the dots represents the p-value of the KS test, with the null hypothesis being that the two distributions being compared are identical. The results indicate that the calculated probability remains relatively stable, irrespective of the chosen threshold. Threshold values ranging from -980 to -900 m appear to yield the most optimal fit for the GPD as indicated by higher p-values and narrower 95% confidence intervals.

5. Discussion

In this work, we presented a novel methodology for estimating mid-air collision risk by combining Monte Carlo simulation and the Extreme Value Theory's Peaks Over Threshold approach. This method offers several key advantages over traditional methods: (i) Reduced computational effort: Our approach decreases the number of Monte Carlo runs, leading to a tenfold reduction in the case study. While this reduction may not seem extraordinary, we would argue that in the presented example, pure Monte Carlo simulation would not have been possible due to the limited amount of historical data that is available. In situations where sufficient historical data is available, a data-driven approach is arguably superior to simulated traffic that often relies on many assumptions and expert judgement. However, independently of how the data is generated, the proposed method makes efficient use of the data. (ii) Quantification of confidence intervals: our methodology allows for the calculation of confidence intervals, an interesting feature not available in traditional collision risk models. In the practical example we showed that the MAC probability can be chosen for different confidence intervals. For example, the nominal estimation is a probability of 1.22×10^{-9} but considering a 95% confidence interval, the obtained probability is below 3.35×10^{-9} leading to a risk of 6.7×10^{-9} (recall from the introduction that one collision counts for two fatal accidents) fatal accidents per departure on runway 16.

For illustration, let's compare the risk of 6.7×10^{-9} fatal accidents per movement to a TLS of 1×10^{-8} fatal accidents per movement. This comparison indicates that implementing a mitigation measure with a cut-off distance of 5NM, rather than 7NM, would yield a risk almost in line with the TLS. It is important to note, however, that mid-air collisions are not the sole cause of fatal accidents. Wake turbulence encounters may also lead to catastrophic incidents, which could be a contributing factor for the current application of a 7NM separation distance instead of 5NM. Furthermore, it is worth mentioning that, due to data quality limitations, certain simplifications have been employed in our analysis (e.g., the assumption of a one minute between take-off clearance and the point when a departure reaches the end of the runway). To operationalize these findings, these simplifications would need to be refined and addressed to ensure a more accurate representation of real-world conditions.

While the proposed methodology offers significant advantages over pure Monte Carlo simulation, it is important to acknowledge its limitations. One must exercise caution in extrapolating the tail too far beyond the observed data. In such a case, the large confidence intervals would provide an indication that the uncertainty is too large to make any reasonable predictions. Additionally, the choice of an appropriate threshold is important, as it can significantly impact the accuracy of the final results. Despite the existence of numerous techniques for threshold selection, each has its own shortcomings and remains susceptible to potential errors. By acknowledging these limitations and exercising caution in the application of the method, our EVT-based approach can still provide valuable insights and contribute meaningfully to the field of mid-air collision risk modelling.

6. Conclusion and outlook

In conclusion, our methodology allows for efficient mid-air collision risk modelling by combining the strengths of Monte Carlo simulation and EVT to optimize the estimation of low probabilities. As a data-driven approach, it enables the accurate computation of risks for complex traffic patterns with relatively fewer assumptions compared to traditional methods. We have demonstrated the effectiveness and applicability of the proposed methodology by examining a real-world scenario, which would have been challenging to assess using pure Monte Carlo simulation. Our findings indicate that this method-

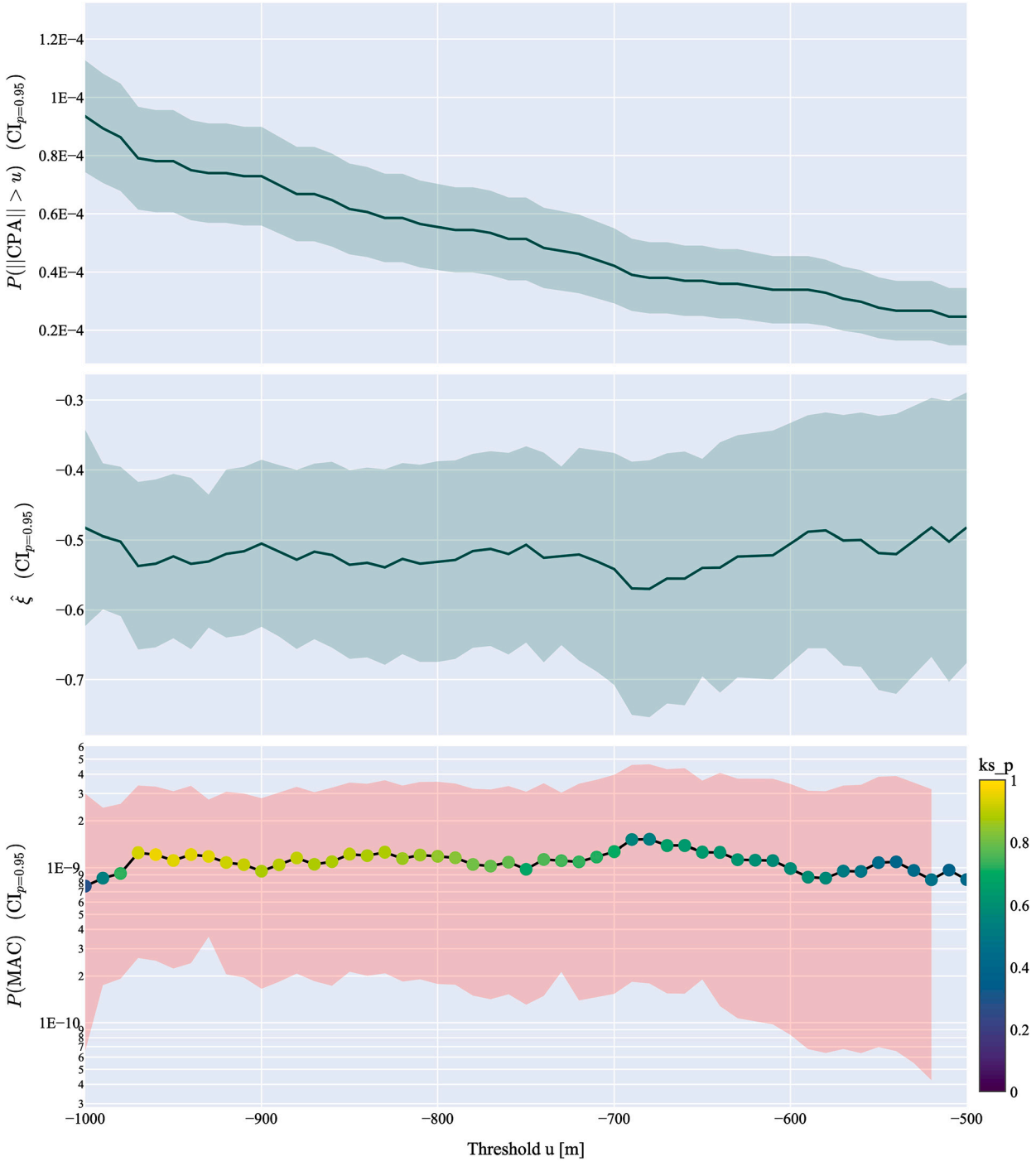


Fig. 11. Variation of the threshold values u and its effect on probabilities. The top plot shows the estimated value $\text{P}(-d_{\text{CPA}} > u)$, the middle plot displays the associated $\hat{\xi}$, and the bottom plot depicts the calculated probability of a collision. For each plot, the respective 95% confidence intervals are represented by the coloured area. The colour of the markers in the bottom plot is linked with the associated Kolmogorov-Smirnov test's p-value of the fitted GPD. A higher p-value indicates a better fit.

ology could potentially improve the efficiency of the use runway 14 in Zurich by refining the current mitigation measures to be less conservative, while still maintaining an appropriate target level of safety. By offering a valuable tool for enhancing the safety and efficiency of air traffic procedures and separation standards, our EVT-based approach contributes meaningfully to the field. Future research could explore the application of this methodology to other types of events, such as loss of separation, airspace infringements, or other safety-critical scenarios in aviation. Moreover, researchers may also investigate opportunities for refining threshold selection techniques or further tailoring the approach to specific contexts, thereby expanding its versatility and applicability across a wide range of aviation safety challenges.

CRediT authorship contribution statement

Benoit Figuet: Conceptualization of this study, Methodology, Software, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization. **Raphael Monstein:** Writing - Original Draft, Writing - Review & Editing. **Manuel Waltert:** Writing - Review & Editing, Project administration, Funding acquisition. **Jérôme Morio:** Writing -

Review & Editing. During the preparation of this work the authors used ChatGPT in order to improve the language quality. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

This work was supported by the Swiss Federal Office of Civil Aviation, grant number BAZL SFLV2018-037.

References

[1] ICAO, Advanced Surface Movement Guidance and Control Systems (A-SMGCS) Manual, ICAO, 2004.
[2] Eurocontrol, Safety minima study: review of existing standards and practices, https://www.eurocontrol.int/sites/default/files/2019-06/src-doc-1-e1.0.pdf, 2000.
[3] ICAO, Regional Supplementary Procedures (Doc 7030), 5 ed., 2008.
[4] X. Lin, N. Fulton, M. Westcott, Target level of safety measures in air transportation – review, validation and recommendations, in: Proceedings of the IASTED International Congress on Advances in Management Science and Risk Assessment (AMSRA 2009), 2009, 662–222.
[5] P.G. Reich, Analysis of long-range air traffic systems: separation standards—i, J. Inst. Navig. 19 (1966) 88–98.
[6] D. Hsu, The evaluation of aircraft collision probabilities at intersecting air routes, J. Inst. Navig. 34 (1981) 78–102.
[7] ICAO, Unified framework for collision risk modelling in support of the manual on airspace planning methodology with further applications, circular 319-an/181 ed, International Civil Aviation Organization, Montreal, Canada, 2008.
[8] ICAO, Doc 9689, manual on airspace planning methodology for the determination of separation minima, International Civil Aviation Organization, Montreal, Canada, 1998.
[9] ICAO, Doc 10116, manual for ats surveillance separation using rcp240 communications, International Civil Aviation Organization, Montreal, Canada, 2008.
[10] M. Mitici, H.A. Blom, Mathematical models for air traffic conflict and collision probability estimation, IEEE Trans. Intell. Transp. Syst. 20 (2018) 1052–1068.
[11] B. Figuet, Collision risk modelling applied to quantitative ifp protection buffer dimensioning, in: ICAO, Separation Airspace Safety Panel, vol. SASP-WGH-36-2022-May-Virtual-IP-15, ICAO, 2022.
[12] B. Figuet, R. Monstein, S. Barry, Data-driven airborne collision risk modelling using a probability density function, in: 15th Air Traffic Management Research and Development Seminar, 2023.
[13] M. Henry, S. Schmitz, N. Revenko, K. Kelbaugh, A Monte Carlo simulation for evaluating airborne collision risk in intersecting runways, AIAA, 2013, pp. 1–21, https://arc.aiaa.org/doi/pdf/10.2514/6.2013-4598, https://arc.aiaa.org/doi/abs/10.2514/6.2013-4598.
[14] H. Fricke, S. Förster, M. Vogel, Using agent-based modeling to determine collision risk in complex tma environments: the turn-onto-ils-final safety case, Aeron. Aerosp. Open Access J. 2 (2018).
[15] M.L. Williams, L.C. Wood, B.J. Nelson, Safety study of closely spaced parallel operations utilizing paired approach, in: 2019 IEEE/AIAA 38th Digital Avionics Systems Conference (DASC), 2019, pp. 1–10.
[16] P. Embrechts, C. Klüppelberg, T. Mikosch, Modelling Extremal Events: for Insurance and Finance, vol. 33, Springer Science & Business Media, 2013.
[17] A.J. McNeil, Estimating the tails of loss severity distributions using extreme value theory, ASTIN Bull. 27 (1997) 117–137.
[18] P. Ruggiero, P.D. Komar, J.C. Allan, Increasing wave heights and extreme value projections: the wave climate of the us Pacific northwest, Coast. Eng. 57 (2010) 539–552.
[19] M. Jacob, C. Neves, D. Vukadinović Greetham, Extreme Value Theory, Springer International Publishing, Cham, 2020, pp. 39–60.
[20] R.A. Fisher, L.H.C. Tippett, Limiting forms of the frequency distribution of the largest or smallest member of a sample, Math. Proc. Camb. Philos. Soc. 24 (1928) 180–190.
[21] B. Gnedenko, Sur la distribution limite du terme maximum d'une serie aleatoire, Ann. Math. 44 (1943) 423–453.
[22] A.A. Balkema, L. De Haan, Residual life time at great age, Ann. Probab. 2 (1974) 792–804.
[23] J. Pickands III, Statistical inference using extreme order statistics, Ann. Stat. (1975) 119–131.
[24] P. Thompson, Y. Cai, D. Reeve, J. Stander, Automated threshold selection methods for extreme wave analysis, Coast. Eng. 56 (2009) 1013–1021.
[25] C. Scarrott, A. MacDonald, A review of extreme value threshold estimation and uncertainty quantification, REVSTAT Stat. J. 10 (2012) 33–60.
[26] J.M. Hoekstra, J. Ellerbroek, Bluesky atc simulator project: an open data and open source approach, in: Proceedings of the 7th International Conference on Research in Air Transportation, vol. 131, FAA/Eurocontrol USA/Europe, 2016, p. 132.
[27] J. Sun, J.M. Hoekstra, J. Ellerbroek, Openap: an open-source aircraft performance model for air transportation studies and simulations, Aerospace 7 (2020) 104.
[28] A. Nuic, D. Poles, V. Mouillet Bada, An advanced aircraft performance model for present and future atm systems, Int. J. Adapt. Control Signal Process. 24 (2010) 850–866.
[29] T. Krauth, A. Lafage, J. Morio, X. Olive, M. Waltert, Deep generative modelling of aircraft trajectories in terminal maneuvering areas, Mach. Learn. Appl. 11 (2023) 100446.
[30] R.B. Millar, Maximum Likelihood Estimation and Inference: with Examples in R, SAS and ADMB, John Wiley & Sons, 2011.
[31] M. Schäfer, M. Strohmeier, V. Lenders, I. Martinovic, M. Wilhelm, Bringing up opensky: a large-scale ads-b sensor network for research, in: IPSN-14 Proceedings of the 13th International Symposium on Information Processing in Sensor Networks, 2014, pp. 83–94.
[32] X. Olive, Traffic, a toolbox for processing and analysing air traffic data, J. Open Sour. Softw. 4 (2019) 1518.
[33] R. Monstein, B. Figuet, T. Krauth, M. Waltert, M. Dettling, Large landing trajectory dataset for go-around analysis, Eng. Proc. 28 (2022) 2.
[34] B.D. Jovanovic, P.S. Levy, A look at the rule of three, Am. Stat. 51 (1997) 137–139.
[35] E.B. Wilson, Probable inference, the law of succession, and statistical inference, J. Am. Stat. Assoc. 22 (1927) 209–212.