29th Congress of the International Council of the Aeronautical Sciences
St. Petersburg, Russia; September 7–12, 2014

INCIDENT PREDICTION USING SUBSET SIMULATION

C. Wang , L. Drees , F. Holzapfel
Institute of Flight System Dynamics (FSD), Technische Universität München (TUM), D-85748 Garching, Germany

Keywords: *accident prediction, flight safety, runway excursion, subset simulation*

Abstract

We obtain statistical distributions of contributing factors, which can potentially lead to an incident from an airline's operational flight data. Then we propagate these distribution through an incident model using subset simulation, which is more efficient than a direct Monte Carlo approach, in order to quantify the incident probability. The method is shown using the incident type of runway overrun. The corresponding overrun model consists of a set of nonlinear differential equations of motion. Some results are computed as examples to show the capabilities of our method.

1 Introduction

Runway excursion is one of the most frequent incidents occurring worldwide [1]. Out of the registered 432 incidents that occurred between 2009 and 2013, 23% of them were identified as runway excursions, making up the largest share. Therefore, many studies focus on determining the typical contributing factors (CF) leading to runway excursions and analyzing their dependencies.

Typical factors that contribute to excursions are high speed deviations from the target approach speed, high tailwinds, landings on a short runway, long landings (touching down late) or wet runways. Usually, studies nowadays analyze the final incident investigation reports in order to determine the main CFs for runway overruns. However, this is unsatisfactory from an airline's perspective because usually the individual flight operation of airlines can vary significantly with respect to aircraft types, procedures, limits, pilot training as well as the route network structure.

Fig. 1 An airline‘s flight operation

The dependencies described in investigation reports might be relevant for that specific airline only. They cannot be reflected in the above mentioned studies that aggregate incidents from airlines worldwide. To overcome this fact, we first derive an overrun model that is based on the dynamics of an aircraft, i.e. a physical approach. Second, we analyze the statistical distribution of the contributing factors by using operational flight data of a single airline. Third, we propagate the statistics of the contributing factors through the overrun model in order to quantify the occurrence probability of an overrun p_{overrun}.

Figure 1 depicts the task that is to be achieved. The goal is to quantify the red area that represents the part of the flight operation in which the considered incident occurs. From a mathematical perspective, the goal is to evaluate a multidimensional integral:

$$p_{\text{overrun}} = \iint_{\text{overrun}} p(\theta)\,\mathrm{d}\theta \qquad (1)$$

where θ refers to the vector of contributing factors with $p(\theta)$ being the probability that a certain combination of θ occurs flight operations. However, as θ is usually of high dimensions, it is not possible to evaluate the integral analytically. Instead, Monte-Carlo simulations can be applied. The incident probability can thus be approximated as follows:

$$p_{\text{overrun}} \approx \frac{1}{N}\sum_{i=1}^{N} I(\theta). \tag{2}$$

I is the indicator function that equals one if an excursion occurred and zero, if the aircraft stops on the runway. By performing a given number N of simulations and simple addition of the number of incidents, one would be able to compute the probability by comparing to the total number of samples. However, for small probabilities, as we are considering in this application, a large sample size is required since N is inverse proportional to the incident probability that is to be computed: $N \sim 1/p_{\text{overrun}}$. Since the probabilities that are to be obtained lie at the order of 10^{-6} or lower, a large amount of samples have to be evaluated, consuming an unacceptable amount of time when using a direct Monte Carlo approach.

2 Subset Simulation

Mathematically speaking, any incident can simply be considered as a failure, i.e. when the load of a system exceeds its capacity. In our case of an runway overrun, the required landing distance of an aircraft would exceed the available landing distance to cause an incident. As this method can be used not only for runway overruns, but also for other incident types, even any failure of a technical system, we will use the term *failure* in general in this section.

In order to reduce the number of samples compared to the direct Monte-Carlo method as mentioned in the previous section, we apply the subset simulation method. The idea is to express the failure domain as a subset of m larger failure domains [2].

$$F_0 \supset F_1 \supset \ldots \supset F_{m-1} \supset F_m = F_{\text{failure}} \tag{3}$$

The probability of the system failure is determined as the product of the conditional probabilities of each subset [2, eq. (2)].

$$\begin{aligned} p_{\text{failure}} &= p(F_m) = p\left(\bigcap_{i=1}^{m} F_i\right) \qquad (4)\\ &= p\left(F_m \middle| \bigcap_{i=1}^{m-1} F_i\right) p\left(\bigcap_{i=1}^{m-1} F_i\right)\\ &= p(F_m|F_{m-1})\, p\left(\bigcap_{i=1}^{m-1} F_i\right) = \ldots\\ &= p(F_1)\prod_{i=1}^{m-1} P(F_{i+1}|F_i) \end{aligned}$$

The failure domains F_i for each subset can be chosen such that the estimated conditional probability $\hat{p}_i$ is equal for all subsets and sufficiently large to be determined using a small number of samples. The last failure domain $F_m = F_{\text{failure}}$ is determined individually but has be equal or greater than the previous conditional probabilities. The final failure probability can then be estimated as:

$$\hat{p}_{\text{failure}} = \prod_{i=1}^{m} \hat{p}_i = \prod_{i=1}^{m} \frac{n_i}{N} = \hat{p}_i^{m-1} \cdot \hat{p}_m, \tag{5}$$

with n_i being the number of samples in each subset that lie in the failure domain and N the total number of samples for each subset. A graphical illustration of the subset simulation is shown in figure 2. One is able to see the samples moving towards the failure domain during the simulation from one subset to the next one.

The first subset is created by using the Monte Carlo method. The input values are sampled from given distributions $q(i)$ that represent the CFs of an airline. The samples in the previous subset θ_i that lie in the failure domain of the previous sample F_i are used to generate the samples for the following subset θ_{i+1}. A Markov Chain using the Metropolis algorithm [3] is applied on every one of the k components of the sample $\theta_i = [\theta_i(1), \theta_i(2), \ldots, \theta_i(k)]$ separately. A proposal distribution p^* is used for the Metropolis algorithm which can be any symmetric distribution:

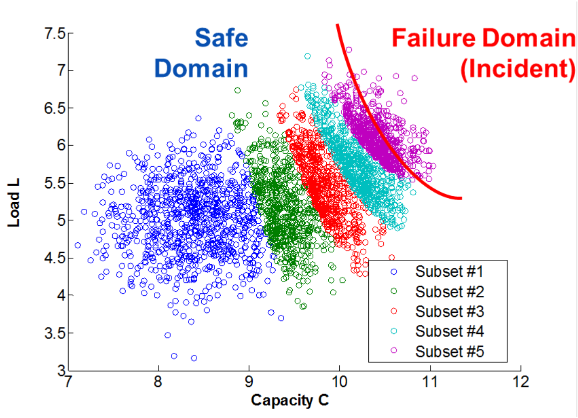

Fig. 2 Concept of Subset Simulation

$p^*(a|b) = p^*(b|a)$, typically a normal or a uniform distribution with a given spread σ_p. The algorithm can be described as follows [2, section 3]:

1. A candidate state ξ_j is generated using the proposal distribution $p^*(\cdot|\theta_i(j))$. The ratio $r = \frac{q(\xi_j)}{q(\theta_i(j))}$ is computed. The candidate sample $\tilde{\theta}_{i+1}$ is set to $\tilde{\theta}_{i+1} = \xi_j$ with the probability $\min(1,r)$ and $\tilde{\theta}_{i+1} = \theta_i$ with the remaining probability.

2. The candidate sample is evaluated using the model. If $\tilde{\theta}_{i+1} \subset F_i$, then it is accepted as a sample for the next subset: $\theta_{i+1} = \tilde{\theta}_{i+1}$. Otherwise set $\theta_{i+1} = \theta_i$.

The Metropolis algorithm ensures that the stationary distribution of the Markov Chain for each component is the input distribution q. For the sampling, each component is transformed to the standard normal space. This enables the same proposal distribution to be applied to every component, regardless of the actual values which can be significantly different if, for example, the air pressure is given in hPa and the approach speed in knots. To summarize, there are four parameters of the subset simulation that can be chosen by the user: The type of proposal distribution (Gaussian or uniform), the standard deviation of proposal distribution σ_p or simply any measure that describes the spread of the distribution, the number of samples per subset N and the conditional probability p_i.

3 Runway Overrun Model

3.1 Physical model

We now return from the description of failures in general to our specific case. As the real-life application, we focus on the runway overrun after landing, which means that the aircraft overshoots the runway. We do not consider lateral motions of the aircraft that could result in a runway veer-off. Therefore, the aircraft can be described as a point mass with several acting forces. The equations describing this point mass can generally be formulated as a set of coupled first-order non-linear differential equations with the vector of states $\mathbf{x}$ and the vector of system inputs $\mathbf{u}$:

$$\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{u})\,. \tag{6}$$

Using Newton's second law of motion, which relates the sum of forces acting on the aircraft $\sum \vec{\mathbf{F}}$ and the time derivative of the linear momentum in an inertial reference frame $(\vec{\mathbf{p}})^I$, we obtain:

$$\sum \vec{\mathbf{F}} = \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^I (\vec{\mathbf{p}})^I\,. \tag{7}$$

The acting forces include aerodynamic forces $\vec{\mathbf{F}}_A^G$, gravitational forces $\vec{\mathbf{F}}_G^G$, propulsion forces $\vec{\mathbf{F}}_P^G$ and landing gear (braking) forces $\vec{\mathbf{F}}_L^G$. The G in the superscript indicates that all forces are acting at the center of gravity. The aerodynamic forces can be described using the aerodynamic coefficients for lift C_L, drag C_D and tranverse force C_Q, along with the aerodynamic speed V_A as well as the air density ρ and the wing reference area S:

$$\vec{\mathbf{F}}_A = \frac{1}{2}\rho V_A^2 S \begin{bmatrix} -C_D \\ C_Q \\ -C_L \end{bmatrix}\,. \tag{8}$$

Propulsion forces $\vec{\mathbf{F}}_P^G$ can be either directly obtained from the manufacturer's datasheet or computed using parameter estimation techniques [4], accounting for air density and temperature. When considering gravitational forces, one has to include the runway slope in both longitudinal and lateral direction. The angles are described by γ and ϕ, respectively:

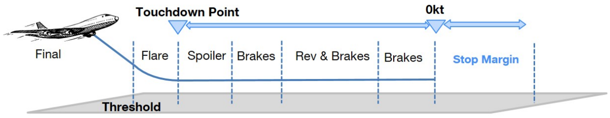

Fig. 3 Calculation of the stop margin

$$\vec{\mathbf{F}}_G = mg \begin{bmatrix} \sin\gamma \\ \cos\gamma\sin\phi \\ \cos\gamma\cos\phi \end{bmatrix}, \tag{9}$$

where m is the aircraft mass and g is the local gravitational constant. Landing gear forces mainly consist of the vertical (normal) forces and the horizontal (braking) forces. The components can be related to each other by using the coefficient of friction μ for the longitudinal and lateral directions x and y:

$$\vec{\mathbf{F}}_L = mg \begin{bmatrix} \mu_x F_N \\ \mu_y F_N \\ -F_N \end{bmatrix}. \tag{10}$$

The sum of forces can therefore be described as:

$$\sum \vec{\mathbf{F}} = \vec{\mathbf{F}}_G^G + \vec{\mathbf{F}}_A^G + \vec{\mathbf{F}}_P^G + \vec{\mathbf{F}}_L^G. \tag{11}$$

Equation 7 can now be numerically integrated once the forces acting on the aircraft are known to obtain the distance required to achieve a speed of zero, i.e. the aircraft reaches full stop after landing. Using the touchdown distances, the runway length, and the deceleration distances, we get values for the stop margin, as seen in figure 3. The input parameters to our simulation model (e.g. headwind, landing weight) are distributed according to the flight operation of the airline that is being considered.

3.2 Operational aspects

The model that was built takes technical failures into account. For example, if the brake system is affected, the forces that can be applied by the brakes are adjusted accordingly. For this case, we make dedicated subset simulations runs since the lack of braking greatly increases the required landing distance and therefore also increases the probability of an overrun. The failure probabilities of each aircraft system can be obtained from maintenance data, for example. In order to compute the probability for an overrun in case of a system failure, e.g. a loss of braking capability, we use conditional probabilities. To obtain the total overrun probability, all relevant failure cases have to be considered. Their results are aggregated in the end.

$$\begin{aligned} &p\,(\text{overrun}) \\ &= p\,(\text{overrun}|\text{no fail}) \cdot p\,(\text{no fail}) \\ &+ p\,(\text{overrun}|\text{brake fail}) \cdot p\,(\text{brake fail}) \\ &+ p\,(\text{overrun}|\text{spoiler fail}) \cdot p\,(\text{spoiler fail}) \\ &+ p\,(\text{overrun}|\text{reverse fail}) \cdot p\,(\text{reverse fail}) \\ &+ \ldots \\ &= p\,(\text{overrun}|\text{no fail}) \cdot p\,(\text{no fail}) \\ &+ \sum_i \underbrace{p\,(\text{overrun}|i\text{ fail})}_{\text{from subset simulation}} \cdot \underbrace{p\,(i\text{ fail})}_{\text{from maintenance}} \end{aligned} \tag{12}$$

By aggregating all failure modes and their respective overrun probabilities, we are able to compute the total overrun probability for a specific aircraft type at a given airport and runway.

An important part of any approach preparation is the determination of the required landing distance by pilots using manuals provided by the aircraft manufacturer and operator. The relevant parameters of the calculation are, for example, weather parameters such as wind and temperature as well as the aircraft weight, but also the status of aircraft systems. If malfunctions occur, particularly for systems that are relevant for the landing performance such as brakes, spoilers or reversers, the required landing distance can increase significantly. The failure of high-lift devices can result in significantly higher approach speeds and subsequently longer landing distances as well. We determine this landing distance required from the operating manual of the specific aircraft for each sample that is generated. If the required distance exceeds the landing distance that is available at the particular runway, we assume that the flight crew would abort the approach and select an alternate longer runway for landing. Thus, we reject this particular sample in this case.

4 Use of operational flight data

The data on which our calculations are based are recorded by the on-board Quick Access Recorder. Flight data recording and Flight Data Monitoring is required for aircraft operators by law in many countries, including the member states of the European Union [5, OPS 1.037]. The data is stored for the entire length of every flight and regularly transferred from the recorder to the database of the airline during routine maintenance checks. The number of recorded parameters vary between different types of aircraft. In a modern fly-by-wire aircraft in which many parameters are already available on the data bus such as the Airbus A340, the number of recorded parameters can be as high as 1600. The recently developed aircraft types are capable to record even more parameters, being more than 3000 on the Airbus A380. They do not only include aircraft states, e.g. speed, altitude or heading, but also input commands such as stick and rudder deflection, thrust lever position or aircraft system parameters including hydraulic and brake system pressure. Environmental parameters are recorded as well, such as wind speed and direction or air density and temperature. The sampling rates of the parameters vary significantly, depending on the specific parameter, but typically range between 0.25 Hz and 8 Hz. Parameters with high dynamics are recorded more frequently than those with little changes over time.

From the recorded flight data, we are able to extract the information required for our approach. For example, if we want to obtain the distribution of the aircraft weight at landing, we have to extract the landing weight from the timeseries data of every flight and collocate them into one histrogram. As probability distributions are required as inputs for our model, we have to fit distributions to the obtained data, as seen in figure 4. It is particular noteworthy that non-Gaussian distributions have to be considered when making the fit. Since we are especially interested in the tail regions, e.g. particular high tailwinds or particular

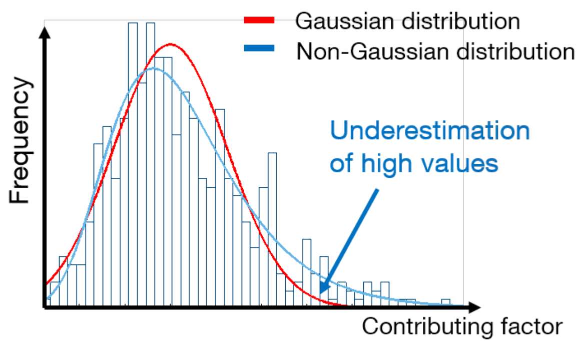

Fig. 4 Fitting of distributions to measured flight data [1]

Table 1 Subset simulation parameters

Type	Gaussian
σ_p	1.2
N	10^4
p_i	0.2

high landing weights, we have to make sure that those regions are well described by the fitted distributions, which usually cannot be achieved with Gaussian distributions, as shown in figure 4

5 Results

As the computation example, we chose a generic runway with a particular high elevation and an available runway length of 2700 meters. The aircraft considered in this case shall be the Airbus A321 with a maximum landing weight (MLW) of 75 tons, as certified. The chosen parameters for the subset simulation can be found in table 1. When looking at the computing time, we can see that 9 subsets are required for the computation, resulting in a total of $9 \cdot 10^5$ model evaluations for each subset simulation while a direct Monte-Carlo simulation has to perform at least 10^6 evaluations, and even significantly more if a certain level of confidence shall be achieved.

A total of 30 subset simulations runs for the identical case were performed to determine the confidence of the estimates. The results can be found in table 2, the 99% confidence intervall is also computed in equation 13 [6, page 391]. In this case, the actual variance is unknown. Therefore, we assume that the variance of the obtained

Table 2 Subset simulation results – MLW of 75 t

Mean value	2.1e-06
Std. deviation	7.5e-07
99 % confidence intervall	
1.8e-06	2.5e-06

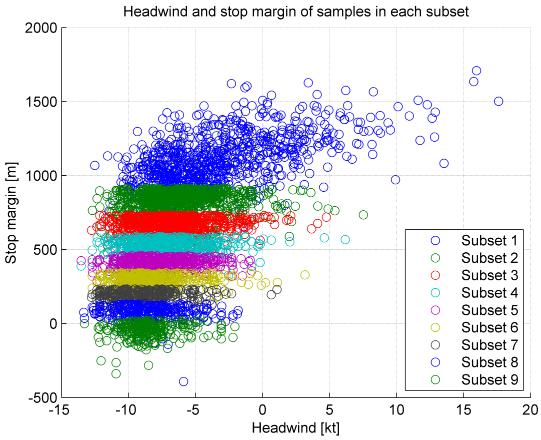

Fig. 5 Headwind versus stop margin – MLW of 75 t

samples S^2 is approximating the actual variance σ^2 while $\bar{X}$ is the mean value of the obtained estimated. n is the number of estimates and $z_{1-\alpha/2}$ is the $(1-\alpha/2)$ quantile of a standard normal distribution. In our specific case, $\alpha = 0.99$ is used.

$$\left[\bar{X} - z_{1-\alpha/2}\frac{S}{\sqrt{n}}, \bar{X} + z_{1-\alpha/2}\frac{S}{\sqrt{n}}\right] \tag{13}$$

When roughly estimating the worldwide runway excursion probability per flight using the numbers of sectors flown by eastern and western-built aircraft [1, page 21] and the number of runway excursions [1, page 30], one would obtain a probability of $6 \cdot 10^{-7}$ for the entire industry in 2013. The mean value we obtained from our approach, accounting for the specific operational aspects of a single airline for this particular airport, appears to be realistic. The standard deviation is about one third of the obtained mean. The 99% confidence interval has a span of $7 \cdot 10^{-7}$, which is also significantly below the mean value. The results can therefore be considered reliable.

Figure 5 shows the samples of the contributing factor headwind plotted against the stop margin for each subset we have obtained. With each subset, the outputs of the runway overrun model

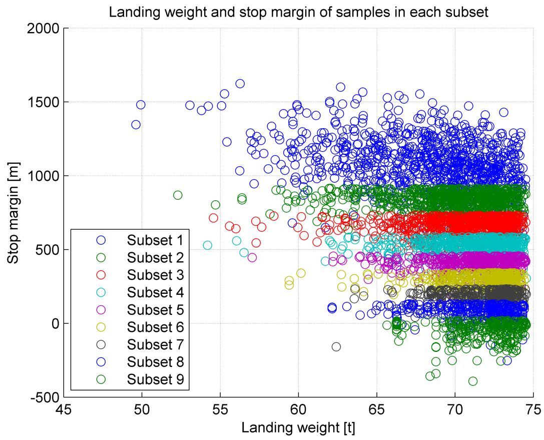

Fig. 6 Landing weight versus stop margin – MLW of 75 t

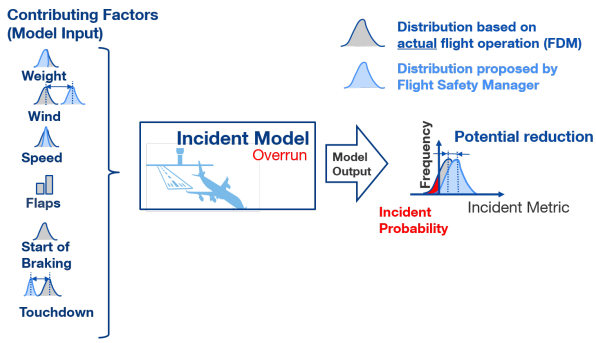

Fig. 7 Impact of changes [1]

move closer to the failure domain, which is expressed by the area with stop margins less than zero. One is able to see nicely the movement of the samples not only to smaller stop margins, but also towards smaller headwinds, i.e. tailwind. We can see that nearly all samples that have negative stop margins, i.e. that result in overruns, have tailwind components as well. The wind can therefore be considered as a significant contributing factor. When looking at the landing weight of the aircraft one would expect that higher weights would lead to longer landing distances. This expectation is met when looking at figure 6. The landing weights are moving towards higher values, however, due to a MLW of 75 tons, the samples cannot exceed this particular limit.

Another benefit of our approach is the possibility to determine the impact of changes within the flight operation and to quatify the effect before those changes are implemented. For exam-

Table 3 Subset simulation results – MLW of 73 t

Mean value	1.4e-06
Std. deviation	5.8e-07
99 % confidence intervall	
1.2e-06	1.7e-06

ple, one measure to decrease the probability of an overrun could be the limitation of a maximum allowable tailwind at landing. Assuming the limits are adhered to by the flight crew, the part of the input distributions for headwind is cut off in which the wind values are stronger than the given allowable tailwind. By propagating the newly obtained distributions through our model, we receive a new value for the probability of an overrun provided that the new distributions correctly describe the changes in flight operation. The concept is shown in figure 7.

In the following, we would like to show the capabilities of our approach when implementing changes in the operation. If the overrun risk is classified as not acceptable by the airline, several measures are possible to reduce the probability, with each of them having both pros and cons. Limiting the maximum allowed tailwind is not always viable since the direction of operation is mainly decided by the airport operator based on a large number of criteria. Another possible method could be the reduction of the MLW with penalties in payload, possibly resulting in financial disadvantages for the airline. In order to decide which measure would be the most effective one, we have to determine the impact of those changes. We want to demonstrated this by reducing the MLW by 2 tons to 73 tons. For this purpose, the landing weight distribution is truncated at 73 tons while the distributions of all other contributing factors remain unchanged. We again propagate the distributions through our model using subset simulation. The results are shown in table 3.

The mean value for the probability of an overrun is reduced by $7 \cdot 10^{-7}$ compared to the scenario with a MLW of 75 tons. The standard deviation is a little higher than one third of the mean. The 99% confidence interval has a range of $5 \cdot 10^{-7}$. We can also see that the confidence intervals of both scenarios do not overlap each other, which means that the reduction of the MLW by 2 tons have a significant impact on the overrun probability. Using those figures, an airline can decide whether a weight penalty with potential financial impact should be accepted to reduce the probability of an overrun in exchange.

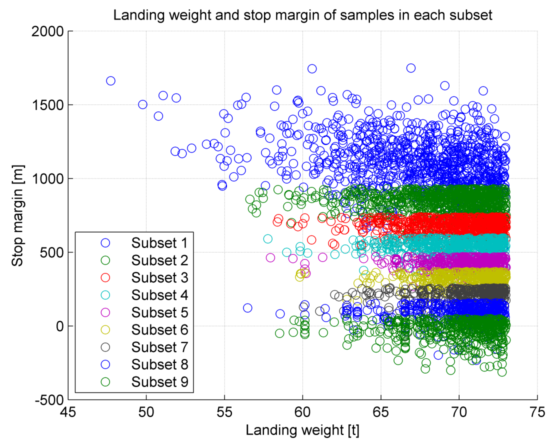

Fig. 8 Landing weight versus Stop Margin – MLW of 73 t

6 Summary and outlook

The subset simulation is a powerful tool to quantify small failure probabilities. Compared to a direct Monte-Carlo approach, the computing time can be significantly reduced. Our application of the subset simulation is the determination of runway overrun probabilities. For this purpose, we developed a physical model that is able to describe the motion of the aircraft during landing using a set of nonlinear differential equations. The model also takes into account specific operational aspects of each airline. The contributing factors of the incident are described as probability distributions that are obtained by fitting distributions to flight operational data which is recorded by the on-board quick access recorder.

The results we obtained are reliable since the standard deviations are well below their corresponding mean values with narrow confidence intervals. The computations are performed using two examples of the same aircraft landing at the same runway with, however, different maximum landing weights being either 73 or 75 tons. The purpose is to demonstrate the ability of our method to quantify the impact of changes in flight operation. A significant difference in the results can be shown if the maximum landing weight is reduces by 2 tons.

Our next steps will include modifications of the sampling method in order to further improve the reliability of the subset simulation algorithm. An approach termed as conditional sampling is already developed, described and applied in [7]. This could be one possibility to further decrease the spread of the results.

Another aspect that is of high significance for airlines are the sensitivities of each contributing factor. This will also be part of our future work. Knowing the sensitivities means knowing the magnitude of impact a change in the contributing factor can cause on the final incident probability. This knowledge will be of great benefit for aircraft operators when measures are to be implemented in order to improve safety. The effects can be obtained in advance.

Finally, other types of incidents can be investigated as well using our methods. The sole work that has to be performed is the buildup of a tailored incident model for each case. Not all incident can be considered, but only those that are physically motivated, which is the majority. Midair collisions, for example, cannot be examined using our approach. Incident types with whom we plan to proceed include other incidents during landing such as general runway excursions, including veer-offs as well as tailstrikes, wingtip strikes and hard landings but also controlled flight into terrain (CFIT) and exceedances of the Mach maximum operating (MMO).

References

[1] IATA. *Safety Report 2013.* International Air Transport Association. 2013

[2] S.-K. Au, J. L. Beck. *Estimation of small failure probabilities in high dimensions by subset simulation.* Probability Engineering Mechanics, 16: 263-277, 2001.

[3] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller und E. Teller. *Equation of State Calculations by Fast Computig Machines.* The Journal of Chemical Physics, Bd. 21, Nr. 6, pp. 1087-1092, 1953.

[4] J. Sembiring, L. Drees, F. Holzapfel. *Extracting Unmeasured Parameters Based on Quick Access Recorder Data Using Parameter-Estimation Method.* AIAA Atmospheric Flight Mechanics Conference, Boston, MA, Aug. 19-22, 2013 AIAA-2013-4848.

[5] European Commission, *COMMISSION REGULATION (EC) No 8/2008 of 11 December 2007 amending Council Regulation (EEC) No 3922/91 as regards common technical requirements and administrative procedures applicable to commercial transportation by aeroplane 2013.* Official Journal of the European Union. 2007.

[6] Fahrmeir, Künstler, Pigeot, Tutz. *Statistik.* 6th edition. Springer. 2007.

[7] I. Papaioannou, K. Zwirglmaier , D. Straub. *Assessment of MCMC algorithms for subset simulation.* Proc. IFIP WG 7.5 Conference on Reliability and Optimization of Structural Systems, Yerevan, Armenia. 2012.

Contact Author Email Address

Chong Wang: chong.wang@tum.de
Ludwig Drees: ludwig.drees@tum.de
Florian Holzapfel: florian.holzapfel@tum.de

Copyright Statement