# Frequency of Rare Event Occurrences 

# ICAO collision risk model for separation minima 

M. Fujita*<br>*ATM Department<br>Electronic Navigation Research Institute (ENRI)<br>Tokyo, Japan<br>m-fujita@enri.go.jp


#### Abstract

Safety is achieved through a continuous risk management process. Risk assessment is an important part of the risk management process. In the risk assessment process, the level of risk in the current/new circumstances is estimated. We introduce the risk assessment methodology developed for the determination of separation minima by the International Civil Aviation Organization (ICAO). It is called the Collision Risk Model (CRM) and estimates the expected number of midair collision occurrences in a given time period or the expected number of fatal accidents per flight hour due to loss of a given separation. Many factors such as aircraft navigation performance and the frequency of position reports are considered to be random variables in CRM. Estimation of the frequency of extremely rare deviations of such random variables is extremely important because a midair collision is considered to be caused by some extremely rare deviations of these random variables. We discuss how to model these random variables so as not to underestimate the frequency of rare event occurrence. We treat the route separation of the parallel route system as an example.


Keywords: separation, route spacing, ICAO collision risk model

## 1. RISK ASSESSMENT

The word "safety" has different connotations depending on one's perspective. For most of the traveling public, it means zero (serious) accidents. However, zero accidents are not achievable. Statistically speaking, even the probability of the reader being injured due to a falling stone is not zero.

ICAO Safety Management Manual ${ }^{[1]}$, which defines a 'Safety Management System' and overviews how to implement it, defines the word "safety" as follows:
Safety is the state in which the risk of harm to persons or of property damage is reduced to, and maintained at or below, an acceptable level through a continuing process of hazard identification and risk management.
The manual remarks that the concept of 'safety' does not mean zero accidents but means that the risk is at or below an acceptable level in the ICAO definition.

The air traffic control environment keeps changing fleet populations and traffic patterns change year by year, and we might implement a new system or a new procedure to meet new users' demands or to mitigate excessive risk. We should continuously monitor current risk level and take remedial action if necessary. Readers should note that safety is achieved through continuous effort.

Risk assessment is an important part of the risk management process. In the risk assessment process, the level of risk in the current/new circumstances is estimated. The results of risk assessment support decision making, indicating whether or not risk mitigation action is required and sometimes proposes which remedial action will be effective.

In the course of wide implementation of a safety management system, the author believes that risk assessment should be conducted much more frequently, and that risk assessment technology should be improved in quality.

## 2. COLLISION RISK MODEL

In air traffic control, spatial and time-dimensional separation between aircraft is established to reduce the risk of aircraft colliding. Separation minima are determined by ICAO and member states to meet specific safety objective.

Two methods of evaluating safety are proposed in [3]. The first method is comparison with a reference system. A new separation minimum is considered to be safe if the collision risk under this separation minimum is proved not to be larger than one of a reference system which has been proved to be safe historically. The second method is the evaluation of risk against a threshold. This threshold is called the target level of safety (TLS). The value of $5.0 \times 10^{-9}[$ fatal accidents per
flight hour] is often utilized. This value was determined based on the historical accident rate.

The value of $5.0 \times 10^{-9}$ [fatal accidents per flight hour] is fairly small and it is impossible to conclude whether the current separation minimum meets the TLS by counting the number of actual collisions. Not only post-implementation safety assessment but also preimplementation safety assessment should be conducted. Of course, counting the number of actual collisions is not applicable for pre-implementation safety assessment. Hence, a mathematical model for estimation of aircraft collision risk has been developed.

ICAO models for collision risk estimation are called collision risk models (CRMs). ICAO CRMs have a long history. The first CRM is found in the papers published by Reich in $1966{ }^{[2]}$. The Reich CRM has been improved and applied for risk estimation of parallel routes (Appendix 4 in [3]) and reduced vertical separation minimum (RVSM) ${ }^{[4]}$. In the Reich CRM, the position error of aircraft is assumed to be timeinvariant. In some cases, we should consider timevariant cases such as collision risk estimation of longitudinal separation in an ADS-C environment ${ }^{[5]}$. At present, CRM is extended to the time-variant case ${ }^{[6]}$.

The proposed fundamental formula in [6] is called the Rice formula. This formula gives the probability that a pair of aircraft will collide. In this formula, aircraft are considered to be points of mass and collision occurs if the relative position of aircraft enters into a volume $\Omega$. In the safety assessment of route spacing, we consider an aircraft as a cuboid with length $=\lambda_{\mathrm{x}}$ (average aircraft length), depth $=\lambda_{y}$ (average wing span) and height $=\lambda_{z}$ (average aircraft height) for simplicity. In this case, $\Omega$ is the cuboid with length $=2 \lambda_{x}$, depth $=$ $2 \lambda_{y}$ and height $=2 \lambda_{z}$ and its center is the origin. When we estimate collision risk of crossing track, an aircraft is assumed to be a cylinder.


Figure 1 Simple Representation of Aircraft Shape

The Rice formula gives the collision probability of a given pair of aircraft during a given time interval as follows:

$$
\begin{equation*}
\operatorname{Pr}\left\{\text { collision during }\left[t_{0}, t_{1}\right]\right\}=\int_{t_{0}}^{t_{1}} \Psi(u) d u \tag{1}
\end{equation*}
$$

Here $\Psi(\mathrm{t})$ is given by

$$
\begin{equation*}
\Psi(t)=\iint_{\partial \Omega}\left(\iiint f_{t}(X, V)(\vec{n} \cdot \vec{V})^{+} d V\right) d S \tag{2}
\end{equation*}
$$

The notation $\partial \Omega$ denotes the boundary of $\Omega$ and $\vec{n}$ is the normal vector of the boundary surface $\partial \Omega$. The notation $f_{t}(X, V)$ is the probability density function on the phase space $(\mathrm{X}, \mathrm{V})$ at the given time instance t . The symbol $(\cdot)^{+}$is the function given by

$$
(x)^{+}=\left\{\begin{array}{cc}
x & x>0  \tag{3}\\
0 & \text { elsewhere }
\end{array}\right.
$$

## 3. EXAMPLE OF CRM: PARALLEL ROUTE

If we apply the Rice formula to the parallel route case and accept some technical assumptions, we can get the following Reich-type formula ${ }^{[6]}$.

$$
\begin{equation*}
N_{a y}=N_{x} P_{y}\left(S_{y}\right) P_{z}(0)\left(1+\frac{\lambda_{x}}{\lambda_{y}} \frac{\overline{\mid \dot{y}} \mid}{\overline{\mid \dot{x}}}+\frac{\lambda_{x}}{\lambda_{z}} \frac{\overline{|\dot{z}|}}{\overline{|\dot{x}|}}\right) \tag{4}
\end{equation*}
$$

The notations utilized in this formula are given in Table 1.

Table 1 CRM Parameters

| Notation | Description |
| :---: | :--- |
| $\mathrm{N}_{\mathrm{ay}}$ | Collision risk defined as the expected <br> number of fatal accidents due to loss of <br> lateral separation per flight hour |
| $\mathrm{N}_{\mathrm{x}}$ | Passing frequency. twice the number of <br> passing events divided by total flight <br> hours |
| $\mathrm{P}_{\mathrm{y}}\left(\mathrm{S}_{\mathrm{y}}\right)$ | Lateral overlap probability. Probability <br> that a pair of aircraft being nominally <br> and laterally separated by $\mathrm{S}_{\mathrm{y}}$ NM <br> overlap laterally. (Figure 2) |
| $\mathrm{P}_{\mathrm{z}}(0)$ | Vertical overlap probability. <br> Probability that a pair of aircraft flying <br> at the same flight level overlap <br> vertically. (Figure 2) |
| $\lambda_{\mathrm{x}}, \lambda_{\mathrm{y}}, \lambda_{\mathrm{z}}$ | Average aircraft length, average wing <br> span and average aircraft height |
| $\overline{\|\dot{x}\|, \overline{\mid \dot{y}}, \overline{\|\dot{z}\|}}$Average longitudinal, lateral and <br> vertical relative velocity of a pair of <br> aircraft which are about to collide. |  |



Figure 2 Lateral and Vertical Overlap

We derive the collision risk formula for a special case under the assumption that all aircraft on the same route at the same flight level cruise at the same speed. We
also assume that aircraft enter into the route system at the same flow rate on each route at each flight level. Consider the case where an N parallel route system is observed for H hours. The notations $\mathrm{n}_{\mathrm{i}, \mathrm{j}}$ and $\mathrm{V}_{\mathrm{i}, \mathrm{j}}$ denote the flow rate (the number of aircraft entering into the route over a unit of time) and the aircraft speed on the route $\mathrm{R}_{\mathrm{i}}$ at the flight level $\mathrm{H}_{\mathrm{j}}$. (The variable $\mathrm{V}_{\mathrm{i}, \mathrm{j}}$ may be a negative value.)

We fix a route $R_{i}$ and flight level $H_{j}$ in consideration. We first find the aircraft density (the number of aircraft on the portion of a given route with a unit length at a given flight level). Consider an aircraft entering into the route system at time 0 . This aircraft proceeds $\left|\mathrm{V}_{\mathrm{i}, \mathrm{j}}\right| \times \mathrm{T}$ in distance for duration T . The number of aircraft entering the route system is $n_{i, j} \times T$. The $n_{i, j} \times T$ aircraft entering into the route system during the duration T lie on the segment with length $=\left|\mathrm{V}_{\mathrm{i}, \mathrm{j}}\right| \times \mathrm{T}$ at time T . Hence the aircraft density is $\left(\mathrm{n}_{\mathrm{i}, \mathrm{j}} \times \mathrm{T}\right) /\left(\left|\mathrm{V}_{\mathrm{i}, \mathrm{j}}\right| \times \mathrm{T}\right)$ $=\mathrm{n}_{\mathrm{i}, \mathrm{j}}\left|\mathrm{V}_{\mathrm{i}, \mathrm{j}}\right|$ and the average longitudinal distance between aircraft is given by $\mid \mathrm{V}_{\mathrm{i}, \mathrm{j}} / \mathrm{n}_{\mathrm{i}, \mathrm{j}}$.

We will find the total number of passing events of aircraft flying on the adjacent routes $R_{i}$ and $R_{i+1}$ at flight level $H_{j}$. Consider an aircraft on the route $R_{i+1}$. The flight distance of the aircraft relative to an aircraft on $R_{i}$ for a sufficiently small duration $\Delta t$ is given by $\left|V_{i+1, j}-V_{i, j}\right| \times \Delta t$. Since the aircraft density is $n_{i, j}\left|V_{i, j}\right|$ on $\mathrm{R}_{\mathrm{i}}$, the expected number of aircraft passing the given individual aircraft on $\mathrm{R}_{\mathrm{i}+1}$ is given by the relative flight distance times aircraft density. Hence, it is

$$
\begin{equation*}
\frac{n_{i, j} \times\left|V_{i+1, j}-V_{i, j}\right|}{\left|V_{i, j}\right|} \Delta t \tag{5}
\end{equation*}
$$

The number of aircraft on $\mathrm{R}_{\mathrm{i}+1}$ is $\mathrm{L} \times$ (density) $=$ $\mathrm{L} \times \mathrm{n}_{\mathrm{i}, \mathrm{j}} /\left|\mathrm{V}_{\mathrm{i}, \mathrm{j}}\right|$, and the total number of passing events of aircraft flying on the adjacent routes $\mathrm{R}_{\mathrm{i}}$ and $\mathrm{R}_{\mathrm{i}+1}$ at flight level $\mathrm{H}_{\mathrm{j}}$ is as follows:

$$
\begin{equation*}
\frac{n_{i, j} \times\left|V_{i+1, j}-V_{i, j}\right|}{\left|V_{i, j}\right|} \Delta t \times \frac{n_{i+1, j} L}{\left|V_{i+1, j}\right|}=\frac{n_{i, j} n_{i+1, j}\left|V_{i+1, j}-V_{i, j}\right|}{\left|V_{i, j}\right|\left|V_{i+1, j}\right|} L \Delta t \tag{6}
\end{equation*}
$$

Hence, the total number of passing events during the observation period H is as follows:

$$
\begin{equation*}
\int_{0}^{H} \frac{n_{i, j} n_{i+1, j}\left|V_{i+1, j}-V_{i, j}\right|}{\left|V_{i, j}\right|\left|V_{i+1, j}\right|} L d t=\frac{n_{i, j} n_{i+1, j}\left|V_{i+1, j}-V_{i, j}\right|}{\left|V_{i, j}\right|\left|V_{i+1, j}\right|} L H \tag{7}
\end{equation*}
$$

On the other hand, the total flight time of aircraft on $\mathrm{R}_{\mathrm{i}}$ at $\mathrm{H}_{\mathrm{j}}$ is given by the average number of aircraft at each instance times observation period, namely, $L \times\left(n_{i, j} /\left|V_{i, j}\right|\right) \times H$. Hence, the passing frequency for the aircraft pairs flying on the adjacent routes $\mathrm{R}_{\mathrm{i}}$ and $\mathrm{R}_{\mathrm{i}+1}$ at flight level $\mathrm{H}_{\mathrm{j}}$ is given by

$$
\begin{align*}
& 2 \times \frac{n_{i, j} n_{i+1, j}\left|V_{i+1, j}-V_{i, j}\right|}{\left|V_{i, j}\right|\left|V_{i+1, j}\right|} L H /\left(\frac{L H n_{i, j}}{\left|V_{i, j}\right|}+\frac{L H n_{i+1, j}}{\left|V_{i+1, j}\right|}\right) .  \tag{8}\\
& \quad=\frac{2 n_{i, j} n_{i+1, j}\left|V_{i+1, j}-V_{i, j}\right|}{n_{i, j}\left|V_{i+1, j}\right|+n_{i+1, j}\left|V_{i, j}\right|}
\end{align*}
$$

The collision risk $\mathrm{N}_{\mathrm{ax}}\left(\mathrm{R}_{\mathrm{i}}, \mathrm{R}_{\mathrm{i}+1} ; \mathrm{H}_{\mathrm{j}}\right)$ for the aircraft pairs flying on the adjacent routes $\mathrm{R}_{\mathrm{i}}$ and $\mathrm{R}_{\mathrm{i}+1}$ at flight level $\mathrm{H}_{\mathrm{j}}$ is given by the following formula by Equations (4) and (8).

$$
\begin{align*}
& N_{a y}\left(R_{i}, R_{i+1} ; H_{j}\right)= \\
& \frac{2 n_{i, j} n_{i+1, j}\left|V_{i+1, j}-V_{i, j}\right|}{n_{i, j}\left|V_{i+1, j}\right|+n_{i+1, j}\left|V_{i, j}\right|} P_{y}\left(S_{i, i+1}\right) P_{z}(0)  \tag{9}\\
& \quad \times\left(1+\frac{v_{y} /\left(2 \lambda_{y}\right)}{\left|V_{i+1, j}-V_{i, j}\right| /\left(2 \lambda_{\mathrm{x}}\right)}+\frac{v_{z} /\left(2 \lambda_{z}\right)}{\left|V_{i+1, j}-V_{i, j}\right| /\left(2 \lambda_{\mathrm{x}}\right)}\right)
\end{align*}
$$

Here $S_{i, i+1}$ denotes the route spacing between $R_{i}$ and $\mathrm{R}_{\mathrm{i}+1}$. Hence the total number of expected collisions in the whole route system during the observation period is given by

$$
\begin{align*}
& \sum_{i=1}^{N-1} \sum_{j=1}^{K} N_{a y}\left(R_{i}, R_{i+1} ; H H_{j}\right)\left(\frac{L H n_{i, j}}{\left|V_{i, j}\right|}+\frac{L H n_{i+1, j}}{\left|V_{i+1, j}\right|}\right) \\
& =L H \sum_{i=1}^{N-1} \sum_{j=1}^{K} \frac{2 n_{i, j} n_{i+1, j}\left|V_{i+1, j}-V_{i, j}\right|}{\left|V_{i+1, j}\right| V_{i, j} \mid} P_{y}\left(S_{i, i+1}\right) P_{z}(0)  \tag{10}\\
& \times\left(1+\frac{v_{y} /\left(2 \lambda_{y}\right)}{\left|V_{i+1, j}-V_{i, j}\right| /\left(2 \lambda_{\mathrm{x}}\right)}+\frac{v_{z} /\left(2 \lambda_{z}\right)}{\left|V_{i+1, j}-V_{i, j}\right| /\left(2 \lambda_{\mathrm{x}}\right)}\right)
\end{align*} .
$$

Since the total flight hours in this route system are

$$
\begin{equation*}
\sum_{i=1}^{N-1} \sum_{j=1}^{K}\left(L \times \frac{n_{i, j}}{\left|V_{i, j}\right|} \times H\right) \tag{11}
\end{equation*}
$$

the collision risk in the whole route system is given as follows:

$$
\begin{align*}
& N_{a y}=\left(\sum_{i=1}^{N-1} \sum_{j=1}^{K} \frac{2 n_{i, j} n_{i+1, j}\left|V_{i+1, j}-V_{i, j}\right|}{\left|V_{i+1, j} \| V_{i, j}\right|} P_{y}\left(S_{i, i+1}\right) P_{z}(0)\right. \\
& \left.\times\left(1+\frac{v_{y} /\left(2 \lambda_{y}\right)}{\left|V_{i+1, j}-V_{i, j}\right| /\left(2 \lambda_{\mathrm{x}}\right)}+\frac{v_{z} /\left(2 \lambda_{z}\right)}{\left|V_{i+1, j}-V_{i, j}\right| /\left(2 \lambda_{\mathrm{x}}\right)}\right)\right)  \tag{12}\\
& \quad /\left(\sum_{i=1}^{N-1} \sum_{j=1}^{K} \frac{n_{i, j}\left|V_{i+1, j}\right|+n_{i+1, j}\left|V_{i, j}\right|}{\left|V_{i+1, j}\right| V_{i, j} \mid}\right)
\end{align*}
$$

It should be remarked that route spacing might be reduced due to the strategic lateral offset procedure (SLOP, lateral offset 1 NM or 2 NM to the right, Figure 3). Aircraft can apply SLOP without informing air traffic control. We cannot know whether aircraft deviate 1 NM intentionally (SLOP) or not from surveillance data. Hence, it is conservative to assume that the values of route spacing of parallel routes are $28 \mathrm{NM}, 30 \mathrm{NM}$ and 26 NM in the case of Figure 3, unless we get reliable data on the proportion of lateral offset application.

| same direction | opposite direction | opposite direction |
| :---: | :---: | :---: |
|  |  |  |
| 1NM 1NM 1NM 1NM | 1NM 1NM 1NM 1NM | 1NM 1NM 1NM 1NM |

Figure 3 Effect of SLOP to Route Spacing (30NM nominal route spacing)

## 4. DISTRIBUTION MODEL

An air traffic controller gets position information through a surveillance system. It is in principal impossible to know the exact position of an aircraft at arbitrary time instance without latency. So aircraft can collide even if the separation among aircraft estimated from the surveillance data is sufficient.

If a potential collision is predicted, the air traffic controller gives some instruction to the relevant aircraft through the communication system. However, if it takes time to transmit the instruction from the ground ATC system to the aircraft, or if the ATC instruction is not transmitted, the aircraft might collide before receiving the instruction.

Aircraft try to fly as initially planned or instructed by ATC. However, the navigation performance of aircraft also has some limitation and a collision due to poor navigation performance is possible.

Air traffic controllers and pilots are human beings and human beings make errors even if they are well disciplined.

Collision risk is determined by the performance of communication, navigation, surveillance, human factors, and so on. All factors are quantified for the estimation of collision risk in CRM. They are often expressed as random variables. What kind of distribution model is appropriate for this purpose? We introduce some methodologies to find a distribution model from a given data set or given requirements.

### 4.1 Estimation from Requirement

The implementation of Performance-Based Navigation (PBN) separation is being promoted. The PBN concept is defined as follows ${ }^{[7]}$ :

The Performance Based Navigation (PBN) concept specifies RNAV system performance requirements in terms of accuracy, integrity, availability, continuity and functionality needed for the proposed operation in the context of a particular airspace concept, when supported by the appropriate navigation infrastructure.

In the PBN concept, RNAV-X and RNP-X aircraft are designed. The expression ' X ' refers to the lateral navigation accuracy of X nautical miles (NM) that is expected to be achieved at least 95 percent of the flight time by the population of aircraft operating within the airspace, route and procedure. RNP-X aircraft have a monitoring and alerting function that ensures that the probability that the total system error (Figure 4) of each aircraft exceeds 2 times X without annunciation is less than $10^{-5}$.


Figure 4 Lateral Navigation Error ${ }^{[7]}$

Let $\mathrm{f}_{\text {TSE }}(\mathrm{u})$ be the probability density function of total system errors (TSEs), then the lateral overlap probability is given by

$$
\begin{align*}
P_{y}\left(S_{y}\right) & =\int_{S_{y}-\lambda_{y}}^{S_{y}+\lambda_{y}} \int_{-\infty}^{+\infty} f_{T S E}(y+u) f_{T S E}(u) d u d y  \tag{13}\\
& \cong 2 \lambda_{y} \int_{-\infty}^{+\infty} f_{T S E}\left(S_{y}+u\right) f_{T S E}(u) d u
\end{align*}
$$

where $S_{y}$ is the route spacing and $\lambda_{y}$ is the average wing span.

Table 2 Convolution for $\mathrm{X}=4$


ICAO SASP (Separation and Airspace Safety Panel) recommends the double exponential (DE) distribution and the Gaussian distribution as the distribution models for RNAV-X and RNP-X aircraft, respectively ${ }^{[8]}$. For the double exponential case, the probability density function is given by

$$
\begin{equation*}
f_{T S E}(y)=\frac{\exp (-|y| / \lambda)}{2 \lambda} \tag{14}
\end{equation*}
$$

where $\lambda=X / 3$. The parameter $\lambda$ is determined to satisfy the following equation:

$$
\begin{equation*}
\int_{-X}^{X} f_{T S E}(x) d x=0.95 \tag{15}
\end{equation*}
$$

On the other hand, for the Gaussian case, the probability density function is given by

$$
\begin{equation*}
f_{T S E}(y)=\frac{\exp \left(-y^{2} /\left(2 \sigma^{2}\right)\right)}{\sqrt{2 \pi} \sigma} \tag{16}
\end{equation*}
$$

where $\sigma=\mathrm{X} / 2.23$. The parameter $\sigma$ is also determined to satisfy the following equation:

$$
\begin{equation*}
\int_{-2 X}^{2 X} f_{T S E}(x) d x=1-10^{-5} \tag{17}
\end{equation*}
$$

Table 2 shows the values of convolution ( $\left.\int_{-\infty}^{+\infty} f_{\text {TSE }}\left(S_{y}+u\right) f_{\text {TSE }}(u) d u\right)$ for $\mathrm{X}=4^{1}$. The left two columns show the convolutions of the probability density function of TSEs satisfying Equation (15). The right two columns show the convolutions of that satisfying Equation (17).

[^0]There is a significant gap in convolution values between the DE and Gaussian cases, and the convolution value decreases more rapidly in the Gaussian case than the DE case. Convolution value (and lateral overlap probability) is more than $10^{7}$ smaller in the Gaussian case than the DE case. If we accept the Gaussian distribution satisfying Equation (17) as the TSE distribution, the collision risk of 30NM route spacing for RNP-4 aircraft is expected to be much less than TLS even in unrealistically heavy traffic cases.

The problem we have is whether the Gaussian distribution is appropriate to use as the TSE distribution model of RNP-X aircraft. We do not have any observation data set on lateral navigation performance of RNP aircraft. It takes a very long time to collect sufficient data to model the type of TSE distribution. At present, we do NOT have any evidence that the TSE distribution does NOT follow the Gaussian distribution, but we do NOT have any evidence that it DOES follow Gaussian distribution.

The choice of distribution type affects the estimation of the risk more than the estimation of flow rate. The estimated collision risk will be approximately halved if the flow rate is mistakenly underestimated by half. (See Equation (12)) On the other hand, if we select the Gaussian distribution in place of the DE distribution, the estimated collision risk becomes more than $10^{7}$ smaller. (See Equation (4) or (12))

Should we accept an assumption in a case where no positive evidence on the validity of the assumption is available, and where this assumption may lead a significant underestimation of the risk? The Gaussian distribution possibly leads a significant underestimation of the risk.

The neighborhood of the peak of a unimodal distribution is called the 'core' of the distribution and the extremities, which are far from the peak, are called 'tails' (See Figure 5).


Figure 5 Core and Tail of Distribution

A collision occurs because of a significant size of TSE. Hence the tail of the TSE distribution dominates the collision risk rather than the core. It is the case for almost all other parameters.

Extreme value theory ${ }^{[9]}$ claims that the conditional probability $\operatorname{Pr}\{\mathrm{Y}<\mathrm{y} \mid \mathrm{Y}>\mathrm{u}\}$ of distributions satisfying certain technical assumptions approximately follows a generalized Pareto distribution (GPD) when $u$ is large enough. More precisely, for any distribution which is in the domain of attraction, $\operatorname{Pr}\{\mathrm{Y}<\mathrm{y} \mid \mathrm{Y}>\mathrm{u}\}$ weakly converges to the GPDs as $u \rightarrow \infty$. The cumulative distribution function of a generalized Pareto distribution is given by

$$
\begin{equation*}
H(y)=1-\left(1+\xi \frac{y}{\sigma}\right)^{-1 / \xi}, \quad 1+\xi y / \sigma>0 \tag{18}
\end{equation*}
$$

When the shape parameter $\xi<0$, the GPDs are Beta distributions. They are exponential distributions and Pareto distributions in the case where $\xi=0$ and $\xi>0$, respectively. Extreme value theory is also introduced in [6].

Since we do not have any information on the type of tail distribution, we should assume that the tail distribution approximately follows a GPD. It should be remarked that GPD is an exponential distribution in cases where the original distribution is Gaussian. Hence, in the TSE case, we should choose the DE distribution rather than the Gaussian distribution for the estimation of collision risk.

DE is not a perfect model either because there is no evidence that TSE distribution follows a DE distribution. However, we cannot make any decision
on the type of distribution if we wait for a sufficient data set. What we should do is to make an assumption on the type of distribution and conduct a safety assessment based on this assumption. This assumption should seem to be sufficiently conservative. After the implementation of a new separation minimum or procedure, we should continue monitoring to check whether a new hazard has been found and whether the assumption made is truly conservative.

### 4.2 Estimation from Data Set

We introduce methodologies to check whether the initial assumption is satisfied based on extreme value theory. Extreme value theory is also applied for the estimation of the longitudinal speed prediction error distribution of ADS-C aircraft ${ }^{[10]}$ in the safety assessment of ADS-C longitudinal separations.

We consider the TSE distribution here. NOPAC (North Pacific) route is one of the most congested route systems in Fukuoka FIR. 50NM route spacing for RNP $10^{2}$ aircraft has been implemented in NOPAC. TSE was assumed to have exponential tails in the safety assessment of this route system.


Figure 6 NOPAC route system

The segment between NUBDA and NANNO is within the radar coverage of Kushiro ARSR by inches and aircraft are expected to follow remote procedure there. Figure 7 shows one-minus-cumulative function of the TSE distribution at the middle point between NUBDA and NANNO on ATS route R220 estimated from Kushiro ARSR data and FDPS data. These data were collected from October 2006 to September 2007. The tail of the TSE distribution seems to be exponential rather than Gaussian.

[^1]

Figure 7 TSE observed at the middle point between NUBDA and NANNO (Oct. 2006 - Sep. 2007)

The author analyzed the tail of Figure 7 by POT (Peak over Threshold). R-package extRemes ${ }^{[11]}$ was utilized for the analysis. First, we set the threshold $u=5$ considering the stability of estimated shape parameter $\xi$ and scale parameter $\sigma$. The number of excesses of thresholds is 44 ( $0.186 \%$ of the whole data set). By the maximum likelihood method (MLM), we found $\xi=$ -0.103 and $95 \%$ confidence interval is [ $-0.518,0.480$ ]. $\sigma=2.92$ and its standard error is 0.765 . Figure 8 and Figure 9 show the QQ-plot and the density plot of this model, respectively. If the empirical distribution follows the distribution model, the dots in QQ-plot are on a straight line.


Figure 8 QQ-plot of TSE distribution


Figure 9 Density plot of TSE distribution

Since the value of $\xi$ is approximately equal to 0 , the TSE distribution seems to have an exponential tail. However, the estimated parameter has a large standard deviation and some dots are somewhat distant from the diagonal line in the QQ-plot diagram because of the small size of data set. Therefore, we cannot get any conclusion from this analysis.

## 5. SUMMARY

ICAO has developed mathematical models which are utilized for quantitative safety assessment in the determination of separation minima. They are called collision risk models (CRMs). We introduce two fundamental collision risk formulae, namely the Reich formula and the Rice formula. The Reich formula is applicable only for time-invariant cases but the Rice formula is an extended version of the Reich formula and is applicable for both time-invariant and timevariant cases. We developed a collision risk formula for a parallel route system in the case where flow rate and aircraft speed on each route at each flight level is constant.

Collision risk is determined by the performance of communication, navigation, surveillance, human factors, and so on. Regardless of the CRM utilized, all factors are quantified for the estimation of collision risk. They are often expressed as random variables. We discussed what kind of distribution model is appropriate for this purpose. We considered a special case - the distribution of total system error (TSE), but the same argument is applicable in many cases. The double exponential (DE) distribution is compared with the Gaussian distribution. We should choose a conservative model so as not to underestimate the risk. The author believes that the DE distribution is better for the model of the TSE distribution. However, the

DE distribution model is not a perfect model, either. We should continue monitoring to check whether a new hazard is found and whether the assumptions on the distribution model are truly conservative enough for risk estimation.

The Generalized Pareto distribution (GPD) was also introduced and applied to the TSE data set. The GPD is a good tool to investigate the shape of tails if a data set of sufficient size is obtained.

## 6. ACKNOWLEDGMENTS

The author would like to thank the Japan Civil Aviation Bureau and Sapporo ACC for the provision of data sets.

## REFERENCES

[1] "Safety Management Manual", ICAO Doc 9859AN/460, 2006
[2] Reich, P. G., "Analysis of long-range air traffic system: separation standards - I, II, III", Journal of the Institute of Navigation, vol. 19 (1966), p.88-98, p.169-186, 331-347.
[3] "Manual on Airspace Planning Methodology for the Determination of Separation Minima", ICAO Doc 9689-AN/953, 1998.
[4] "Manual on Implementation of a 300 m ( 1000 ft ) Vertical Separation Minimum Between FL 290 and FL 410 Inclusive", ICAO Doc. 9574-AN/934 (2 $2^{\text {nd }}$ ed.), 2002.
[5] Fujita, M, "Safety Assessment prior to 30NM Longitudinal Separation Minimum under ADS-C Environment", ICAO SASP/WG/WHL/13 IP/8, Montreal, May 2008.
[6] "A unified framework for collision risk modelling in support of the manual on airspace planning methodology for further application" ICAO Circ 319-AN/181, advance edition.
[7] "Performance-Based Navigation Manual", ICAO Doc. 9613-AN/937 ( ${ }^{\text {rd }}$ ed.), 2008.
[8] "Report of the Mathematicians Sub Group meeting at the SASP WG/WHL/10 meeting", ICAO SASP/WG/WHL/10, Gold Coast, November 2006.
[9] Coles, S, "An Introduction to Statistical Modeling of Extreme Values", Springer, 2001.
[10] Fujita, M, "Distribution of Longitudinal Speed Prediction Error of ADS-C System", $3^{\text {rd }}$ International Conference on Research in Air Transportation (ICRAT2008), p.143-150.
[11] http://www.assessment.ucar.edu/toolkit/

## Copyright Statement

The author confirms that he, and/or his company or institution, hold copyright of all original material included in their paper. He also confirms he has obtained permission, from the copyright holder of any third party material included in their paper, to publish it as part of their paper. The author grants full permission for the publication and distribution of his paper as part of the EIWAC 2009 proceedings or as individual off-prints from the proceedings.


[^0]:    ${ }^{1}$ RNP4 is a required navigation performance for 30 NM route spacing in oceanic/remote airspace.

[^1]:    ${ }^{2}$ Note that RNAV10 in oceanic/remote airspace in the PBN definition is usually called RNP10 for historical reasons.

