Article citation information:

Telesfor Marek MARKIEWICZ¹, Piotr MAKOWSKI²

**SELECTED PROBLEMS OF RISK EVALUATION OF CHANGE IN CERTIFIED CIVIL AVIATION ORGANISATIONS**

**Summary.** This article presents some of the generalised results of this research conducted in the form of a diagnostic survey among Safety Managers and a review of safety risk management procedures, including the rules of change risk evaluation, applied in selected civil aviation organisations subject to the survey. This publication aimed to put them under the discussion of experts interested in the issues. Important problems resulting from the necessity to meet general guidelines standardising processes of risk management, formulated in international law regulations [2], as well as in advisory materials of the International Civil Aviation Organisation [6-8], and others [3, 4], including national ones [11], in confrontation with the need for their practical application, were described. Concerning the general rule of risk evaluation described there, inconsistencies in interpretation of the scale of probability and frequency indices and their influence on the evaluation context were pointed out. A model, considering the time aspect in Event Tree Analysis - ETA procedures based on [5, 9], was proposed. Perceived threats to the results of change risk evaluation caused by applying the "the worst foreseeable situation" principle were described. A different approach in the evaluation of change risk based on a comprehensive assessment of the impact of all identified risks on

¹ War Studies University, Chrusciela 103 street, Warsaw 01-910, Poland. Email: m.t.markiewicz@akademia.mil.pl. ORCID: 0000-0001-6085-5906

² War Studies University, Chrusciela 103 street, Warsaw 01-910, Poland. Email: p.makowski@akademia.mil.pl. ORCID: 0000-0002-3045-3495
the defined effect was also proposed. Proposals were made to eliminate the disadvantage resulting from the definition of the probability index value – "extremely improbate".

**Keywords:** safety risk, risk of change, aviation safety management system

1. **INTRODUCTION**

The provision of services by certified civil aviation organisations is associated with the need for appropriate safety management. Air accidents and serious incidents, in addition to direct severe losses, carry the threat of indirect losses, including loss of reputation and anticipated profits. Safety management, including safety risk management (SRM), has become not only an objective necessity, but also an obligatory requirement imposed by the legal regulations of many countries [2], and enforced by their aviation authorities [11]. This is a consequence of the provisions of Annex 19 of the Chicago Convention [2], according to which civil aviation organisations must implement a safety management system (SMS). More important, lectures defining the scope of SMS tasks are successive editions of Doc 9859 Safety Management Manual (SMM) issued by ICAO [6-8]. The system should also include procedures for the assessment of risk associated with the implementation of changes in the organisation [2]. Therefore, the risk should be classified according to the Risk Classification Scheme (RCS) adopted by the organisation.

Safety risk evaluation in the light of SMS adopted SRM standards completes the stage of risk assessment and aims to express the degree of tolerance towards risk. The authors' participation in the review of the safety assessment process in Polish Air Navigation Services Agency (PANSA) and pilot studies of the area in question in other organisations allowed identifying some problems of SRM tasks implementation concerning the introduced changes in organisations, resulting from the adopted procedures. The authors' participation in the research team reviewing safety and risk management procedures for PANSA and the results of the diagnostic survey conducted in other civil aviation organisations (expert interviews with selected Safety Managers) allowed collecting many observations of existing SRM problem situations. Among many methodical problems accompanying SRM, identification of difficulties in the risk evaluation phase was a certain surprise. The observations made and recommendations proposed in this article - as results of certain generalisations - were considered important. There was an argument to submit them to a wider judgement of the environment concerned with the creation of SMS in civil aviation organisations in the form of this article.

2. **FORM AND RULE OF SAFETY RISK EVALUATION**

2.1. Risk matrix

The definition of safety risk in the fourth edition SMM is regulatory in nature. Safety risk is defined there as: "The predicted probability and severity of the consequences or outcomes of a hazard" [7, p. (viii)]. It follows that it is, among other things, an indicator of the significance of

---

3 ICAO - The International Civil Aviation Organization, founded in 1944, has 193 Member States

4 According to SMM Doc 9859, the SRM standard distinguishes between the stages of "hazard identification, risk assessment (risk analysis and evaluation) and the stage of safety risk mitigation and acceptance" [7, 2 - p. 10]
Selected problems of risk evaluation of change in certified civil aviation organisations

the anticipated, possible, negative consequences of a change. Risk so defined is identified in a two-criteria space – probability and severity of consequences and can not have other indicators\(^5\). Hence, risk matrices have been adopted as forms of risk evaluation (Figure 1). In the quoted definition, there is no concept of frequency, but in the definition of the risk matrix scale, one can find references to the category of frequency, which is a formal inconsistency.

<table>
<thead>
<tr>
<th>Safety Risk</th>
<th>Catastrophic A</th>
<th>Hazardous B</th>
<th>Major C</th>
<th>Minor D</th>
<th>Negligible E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequent</td>
<td>5A</td>
<td>5B</td>
<td>5C</td>
<td>5D</td>
<td>5E</td>
</tr>
<tr>
<td>Occasional</td>
<td>4A</td>
<td>4B</td>
<td>4C</td>
<td>4D</td>
<td>4E</td>
</tr>
<tr>
<td>Remote</td>
<td>3A</td>
<td>3B</td>
<td>3C</td>
<td>3D</td>
<td>3E</td>
</tr>
<tr>
<td>Improbable</td>
<td>2A</td>
<td>2B</td>
<td>2C</td>
<td>2D</td>
<td>2E</td>
</tr>
<tr>
<td>Extremely improbable</td>
<td>1A</td>
<td>1B</td>
<td>1C</td>
<td>1D</td>
<td>1E</td>
</tr>
</tbody>
</table>

In the RCS, the scales of consequence severity and probability/frequency indices are imposed by regulations [11]. The severity of consequence index scale \{A; B; C; D; E\} is ordinal in nature and the definitions of these indexes correspond in practice to the categories of aviation incidents, classified respectively into: crashes and severe accidents, accidents, major incidents, incidents and low consequence incidents. Hence, these scales do not have a common measure; moreover, the first two indices, A and B, are for Harmful Effect (HE) events associated with personal losses. The probability/frequency indices \{1; 2; 3; 4; 5\} have a common measure. However, the frequency/probability interval for index 1 (extremely improbable) is unilaterally constrained, which creates significant interpretation problems. In the risk assessment matrix (risk tolerance matrix) of Figure 1, a three-tier scale of risk ratings is adopted: "risk intolerable" for risk indices: \{5A, 5B, 5C, 4A, 4B\}; "risk tolerable" for risk indices: \{5D, 5E, 4C, 4D, 4E, 3B, 3C, 3D, 2A, 2B, 2C, 1A\}; "risk acceptable" – for the remaining risk indices. This procedure is the recommended risk evaluation rule in the SMM.

In the proposed evaluation rule, the classification of safety risk as tolerable for risk index \{1A\} is questionable. It can be observed that any flight operation with a crew on board may be at risk of a catastrophe (A), although the probability of its occurrence may be extremely small \{1\}. If the above sentence is true, then its consequence is that the risk of any flight operation should be qualified at least as tolerable, and any actions mitigating this risk does not change its assessment, because there is no frequency index scale to express the effects of these actions. In turn, the classification of risk as "risk tolerable" usually forces, in light of change management procedures, the search for mitigation measures - which, colloquially speaking, "loops the process". Furthermore, the risk matrix used in the evaluation of risks associated with hazards that result in a distribution of aviation events other than binomial loses the property of transparency because of the classification of such risks.

\(^5\) In order not to formally restrict the possibility to estimate risk using other measures and indicators, it is proposed to adopt a general definition of risk as follows: a risk is a feature of a situation in which a projected random development may bring about only negative, negative or positive effects, and the probability distribution of these effects is known, identified with acceptable accuracy [10, p.12]
2.2. Frequency or probability

Both frequency and probability allow expressing the attitude towards the consequences of a given threat generated by a change. However, these categories cannot be equated. Frequency as a physical quantity has units of measurement, unlike probability. Frequency is expressed here as the number of aviation events related to time (for example, flight hours). The practice of safety monitoring in aviation organisations has established some retrospective safety indicators called Safety Performance Indicators - SPIs, referring to frequency and intensity, defined as the number of aviation events of a given type to, for example, 1E5 flight operations [11].

The SMM did not define the values of these interval boundaries. Probability (frequency) was described verbally using an ordinal qualitative scale based on the following linguistic variables: frequent - likely to occur many times (occurred frequently) - index {5}; occasional - likely to occur occasionally (occurred not very often) - index {4}; remote - likely not to occur but possible (occurred rarely) - index {3}; improbable - very unlikely to occur (occurrence is unknown) - index {2}; extremely improbable - almost inconceivable that it could ever occur - index {1} [6, p. 38].

The operation of frequency in risk evaluation makes the evaluation horizon indeterminate. The resulting risk level is independent of the time/number of operations as long as the frequency is stationary. On the other hand, it is obvious that if the frequency is stationary, the probability of occurrence of aviation events increases non-linearly with the increase of time/number of operations. Hence, the level of risk here depends on the assessment horizon. The relationship between probability and time is linear to a small extent for very small values of frequency. As a first approximation, one can assume mutual independence of aerospace events in the sense of probability calculus, and for very small probabilities ignore the phenomenon of their accumulation, which allows one to assume that the assumptions of the Markov process are satisfied. Markov process is a process with discrete states, if they are convertible, it consists of a random selection of the moment, jumping from one state to another, while the probability of changing the state does not depend on the history [5, 9]. Then the probability of remaining in the state of non-occurrence of the air event of interest, \( P_0(t) \), for intensity \( \lambda \) satisfies Andrey's differential equation N. Kolmogorov:

\[
P_0'(t) = -\lambda P_0(t).
\]

Solving equation (1) for initial conditions \( P_0(0) = 1 \) and considering that \( P_0(t) \) is the probability opposite to the probability of occurrence of one aerial event \( P_{k=1}(t) \), we obtain, after elementary transformations, a relation allowing to calculate the mathematical probability of occurrence of an event (k=1) depending on time \( P_{k=1}(t) \):

\[
P_{k=1}(t) = 1 - e^{-\lambda t}.
\]

Relation (2) allows estimating for which range of values: \( \lambda, t \) we are dealing with an approximate linear relationship between probability and time (for example, developing Euler number into Maclaurin series) [9].
Nowadays, the degree of computational complexity has lost its importance, hence, for example, various event models can be analysed in this approach. A common analytical tool for hazard identification in civil aviation organisations is the Bow-Tie diagram\(^6\), the right side of which is commonly known as the Event Tree Analysis - ETA. The result of this analysis is the identified various consequences resulting from a given hazard. When monitoring the results of a change, the presence of a hazard is indicated by an observed symptom. This can be a top event, a trigger event, or any other low consequence event for which a Safety Monitoring Indicator (SPI) has been established. A reactive safety monitoring indicator provides an estimate of the average frequency of a given low consequence event over an assumed evaluation horizon. In this situation, the graph in Figure 2 can be used for safety analysis.

State "1" corresponds to the safe operation of the entity in the operational environment, and \( P_1(t) \) is the probability of being in this state. State "2" corresponds to the occurrence of a Top Event, which in a general case can transform into events with different consequences - from "3"; "4"; to "m".

State changes are characterised by known transition frequencies, respectively: \( \{ \lambda_{12}; \lambda_{23}; \lambda_{24}; \ldots \lambda_{2m} \} \).

![Fig. 2. Example graph of possible consequences of a top event for a time-sensitive ETA interpretation](Source: Authors' elaboration)

The graph defined in this way can be described by a system of differential equations (3), forming a dynamic model [9], allowing, for example, to interpret the ETA diagram with respect to time:

\[
\begin{align*}
P_1'(t) &= -\lambda_{12}P_1(t) \\
P_2'(t) &= \lambda_{12}P_1(t) - (\lambda_{23} + \lambda_{24} + \ldots + \lambda_{2m})P_2(t) \\
P_3'(t) &= \lambda_{23}P_1(t) \\
P_4'(t) &= \lambda_{24}P_1(t) \\
&\vdots \\
P_m'(t) &= \lambda_{2m}P_1(t)
\end{align*}
\]

(3)

\(^6\) The Bow-Tie diagram is one of the most important tools supporting hazard identification and is recommended for use by the Civil Aviation Authority (CAA) [11, p. 53-54]
Solving the system of equations for $P1(0)=1$ and zero initial conditions for the other probabilities allows the evaluation of critical time periods for acceptable risks associated with all identified consequences. The frequency of a phenomenon is easier to estimate from empirical data than the probability of occurrence. Frequency can be interpreted as the average value of the probability distribution of occurrence of aviation events with specified consequences, and thus, has the advantage over probability that, as a statistical measure, it meets the postulate of additivity. Experts are more willing to estimate the frequency, for example, for hazards caused by human factors. Operating with probability in estimating the risk of failure of technical systems, for example, CNS (Communications, Navigation, Surveillance), results from the character of methods used in the reliability theory. Comparison of these two categories of risk in a common scale of risk assessment requires a clear definition of the assessment context, including a common time horizon for assessments.

2.3. The principle of “the worst foreseeable situation”

For the safety risk evaluation procedure, it is important what the results of the risk analysis refer to and how they are relevant to the purpose of the evaluation. Here, it is important whether the sub-assessments are not redundant and whether information is not lost in the procedures. In the course of the analysis of the review of SRM procedures in the selected organisations, some danger was perceived in the obligatory application of the "the worst foreseeable situation" principle.

The recommendation of this principle can be found in all editions of the SMM [6-8]. It was also part of the definition of safety risk\(^7\). The essence of the principle is to consider in risk analysis, the critical cases for which the risk or consequences are the worst of those considered. Repeated application of this principle in different moments of the risk assessment may cause loss of information and underestimation of the result. The notion of 'the worst foreseeable situation' in conjunction with the frequency/probability index extremely improbable is a cause of obvious discomfort to those formulating the expert risk assessment. This is because they have to decide for themselves from which frequency value there is a real possibility of a certain risk effect occurring. It is difficult to find an explanation of this in the SMM. The aviation authorities of ICAO Member States are largely silent on this issue, although the attempt made by the Nepalese aviation authorities is noteworthy. In light of this interpretation: "An occurrence is considered foreseeable if any reasonable person could have expected the kind of occurrence to have happened under the same circumstances. Identification of every conceivable or theoretically possible hazard is not possible" [4, p. 44]. However, this editorial can hardly be considered as a satisfactory explanation.

In this situation, a solution is imposed by changing the risk evaluation rule for the matrix element with index \{1A\}. For this risk index, it is worth considering assigning it a risk level - "risk acceptable", which would mean at the same time that in this case, the rule "the worst foreseeable situation" does not apply. This also removes the disadvantages described in subsection 2.1.

The repeated use of the "worst foreseeable situation" principle is evidenced by the content of the hazard record forms used in civil aviation organisations. Figure 3 shows an example of the form of the hazard record heading from the actual risk assessment documentation and indicates the assessment situations where this principle is considered (Option 1 and Option 2).

---

\(^7\) For example, in SMM (2\textsuperscript{nd} ed.), safety risk is defined as an assessment, expressed in terms of predicted probability and severity, of the consequences of a hazard, taking as reference the worst foreseeable situation [8]
If it is assumed that in a general case a change generates several hazards, and each hazard generates several consequences of varying severity, then in assessing the risk from a hazard, in the light of the form of tools used (Figure 3), the worst consequence is considered (Option 1, Figure 3), which in turn, in assessing its frequency of occurrence may have a more favourable risk index than the risk index for another consequence that occurs more frequently. In contrast, applying the rule to the worst risk index (Option 2) resulting from a hazard limits the evaluations to that one hazard. While the anticipated safety state also depends on whether there are other hazards with sub-critical risk in the operational environment in addition to the critical risk associated with a given hazard and what the magnitude of this phenomenon is.

3. IMPLICATIONS OF INTERPRETING THE CONCEPT OF CHANGE RISK FOR ITS EVALUATION

If change-related safety risk is to be one of the criteria for making decisions in change management, then its essence must be clearly defined. Is it the operational risk component of the change, understood as the difference between the operational risk assessed before the change and the operational risk after the change or is it the operational risk assessed after the change.

It is worth noting that risk is usually associated, in terms of the area of analysis, with the type of activity, (for example, hazardous materials transportation operations, and here the change), the given hazard or its source, and the potential outcome of taking the risk. Further, it is useful to make this explicit. Establishing this element of the risk assessment context depends
on the nature of the change and is essential to interpreting the subject of the assessment unambiguously and clarifying expert judgement about the risk.

Thus, if the risk of a change is classified based on "the worst foreseeable situation" concerning a critical hazard (Option 2, Figure 3), it is difficult to speak of a comprehensive risk assessment. Given that, in a general case, a given change may generate several hazards, each with several effects of varying severity and frequency, there may be a situation where the effect of the safety risk classification on this change requires a comprehensive assessment. In its simplest interpretation, 'comprehensive assessment' means considering all the most important determinants of the assessment. Thus, a safety assessment of a change (or operational safety) may be conducted based on an expert judgement of the risk matrix's compilation of the results of the classification of the individual risks corresponding to the identified hazards, for example, for 'the worst foreseeable situation', if there are only a few of them. With a larger number of hazards and corresponding risks, the expert may lose control of the assessment. The search for a comprehensive risk evaluation index for different hazards and corresponding different distributions of effects (treated as a random variable), is hampered by the lack of a common additive measure of effects, including the occurrence of Harmful Effect events.

In such risk evaluation situations, there is a need to establish an evaluation method that considers the possibility of aggregating the effects resulting from individual hazards with an equal severity index. This is particularly important in situations where a change results in independent hazards that may have similar consequences leading to aerial incidents or degradation of system functions at multiple locations with similar effects. In this situation, a starting point would be an effect severity index, for which the summed frequency resulting from the impact of all hazards should be determined. The property of additivity of the frequency as a predicted average value allows, as a first approximation, to sum up the frequencies of occurrence of a given effect with a given index, estimated independently for individual hazards. The resulting frequency index allows the determination of a new risk index, which will relate to the aggregated effect described by the given index. Generally, the impact of all hazards can be considered in this way for the individual severity indices of the risk matrix. It can be shown that considering the aggregation of effects resulting from individual hazards can significantly change the outcome of the risk classification.

The difference in the results and benefits of risk evaluation using the recommended approach (Option 3) regarding the one currently used (Option 2) is justified by the following example.

Example.

Assume that a change generates two independent hazards, H1 and H2. Threat H1 has been assessed by applying the principle of "the worst foreseeable situation" as critical (according to Option 2). Adopting the index designations as for the risk matrix of Figure 1, for H1, the effect was estimated with the index {B-Hazardous}, and the frequency of its occurrence was estimated to be about once every year and a half, which on the adopted scale corresponds to the index {3-Remote}.

For hazard H2, the worst effect was assessed, corresponding to the index {B-Hazardous} and the frequency of its occurrence was estimated to be about once every two years - which also corresponds to the frequency index {2-Remote}.

---

8 In the case of the category of probabilities, the known relation for the probability of a sum of events can be used. In the case of very small probabilities, one can approximately estimate the probability of a sum of events by summing the probabilities of these events.
Whereby, for the civil aviation organisation discussed in the example, the scales of the frequency indices of interest were established as follows:
- Remote: this effect will not occur more than once per year, but less than once every five years;
- Occasional: this effect will not occur more than once every three months, but at least once per year.

The risk evaluation may be performed according to the procedure used (Option 2) or proposed (Option 3) as follows:
- Option 2 - without consideration of the aggregation of effects, applying the principle of "the worst foreseeable situation" to the critical hazard, in which case, the risk index will take the value as for H1 {3B}, thus defining the outcome of the risk evaluation as "risk tolerable";
- Option 3 (recommended) - considering the aggregation of effects, the risk of an effect with an index {B-Hazardous} will have a frequency index {4-Occasional}, because it has been estimated that the event corresponds to a frequency resulting from the summation of the frequencies of the two hazards (H1 and H2). Thus, it may occur more than once a year, but less than once every three months - which corresponds to a risk index {4B} and an evaluation result: "risk intolerable".

The situation from the example is illustrated in Figure 4.
Furthermore, it is worth noting that even in expert assessments, when estimating frequencies, it is good to operate the forecasts abstracting from the frequency indices and defining the limits of the corresponding value ranges. In the case of applying the approach: Option 3, this can be of great importance in the summation of the frequency components for a given effect generated by different hazards. Therefore, there is a need to consider this in the hazard/risk register forms (Figure 3).

Overall, without applying the "worst foreseeable situation" risk elimination principle, the evaluation of the risk of change must consider all the risks caused by the change. These may generate different impacts classified by all the indices of the severity scale of the risk matrix.
In this assessment situation, we may obtain a large number of change risk component indices exceeding the number of risk matrix elements. This may result in a loss of assessment control. As a result of the proposed approach, in a change risk evaluation after considering the aggregated impacts of all hazards on the expected effects classified according to their indices, we may obtain a risk matrix with one change risk component evaluation in each column, which reduces the number and facilitates the assessment of the riskiness of the new situation following the change.

4. CONCLUSIONS

In certified aviation organisations, given the high safety status of their services and the nature of randomness, empirical verification of the accuracy of safety risk assessments is practically extremely difficult. Hence, the proposals described in this article concerning the rules for the evaluation of the risk of change can be evaluated based on arguments resulting from the justification. The most important generalisations of the content of this article are formulated in the following conclusions:

- when evaluating the risk of a hazard, applying the "the worst foreseeable situation" principle only to the effect severity aspect may underestimate the risk level as a result of its evaluation;
- when a change generates more than one hazard, the application of "the worst foreseeable situation" in selecting the critical risk index value from among the risk indices associated with the hazards (Option 1) may underestimate the risk level;
- the repeated application of "the worst foreseeable situation" in different phases of the risk analysis and evaluation may lead to loss of information and underestimation of the risk level in the risk assessment;
- for the risk associated with the change, treated as a comprehensive criterion for the assessment of the level of safety for the change, due to the limitations of the severity of the effects of the risk matrix, it is proposed to determine an index of the risk in question ensuring the comprehensiveness of the evaluation in the area of effects with the same severity index, and then choose the critical value of the risk index among the risk indices associated with the effects, applying the principle of 'the worst foreseeable situation' (Option 3);
- during the risk analysis, the predicted frequency of the effect should be estimated (not the predicted frequency index), and the frequency index should be derived from this;
- the assignment in the evaluation rule of the risk index \{1A\} (Figure 1) - the risk level: "risk acceptable" removes the interpretation inconvenience of the principle "the worst foreseeable situation" concerning the effects for which the frequency of occurrence is defined vaguely as: "extremely improbate".

If the presented results of the analysis of procedures and rules of risk evaluation would be considered and positively evaluated by those interested in this issue, it would be a source of satisfaction to the authors.
Selected problems of risk evaluation of change in certified civil aviation organisations

References


Received 04.10.2021; accepted in revised form 05.12.2021

Scientific Journal of Silesian University of Technology. Series Transport is licensed under a Creative Commons Attribution 4.0 International License.