



## Executive summary

# Study of the Quality of Safety Assessment Methodology in Air Transport

### **Problem area**

Air traffic is expected to double within the next 20 years, which requires large changes in airspace structure and organisation. The aim of this paper is to determine what is required for a safety validation approach to be appropriate for handling such large changes, and to test if existing approaches already satisfy this requirement.

### **Description of work**

A consolidated set of indicators has been developed, each of which describes one required aspect of a safety validation approach for large changes, and the complete set describes all aspects. Subsequently, the indicators are used to evaluate current safety assessment approaches in air transport.

### **Results and conclusions**

The consolidated set includes indicators related to the scoping of safety validation, to coverage of certain aspects of the operational concept, to risk assessment, to feedback to concept of operations development, to organisation of safety assessment, and to supporting decision and policy makers. The evaluation of current safety assessment approaches in air transport against the indicators shows needs for development of approaches towards safety assessment of large changes in air transport operations.

### **Applicability**

The indicators can be used as guideline to verify if an existing safety validation covers all aspects that are required for safety validation of large changes in air transport operations.

### **Report no.**

NLR-TP-2007-978

### **Author(s)**

M.H.C. Everdij  
H.A.P. Blom

### **Report classification**

UNCLASSIFIED

### **Date**

January 2008

### **Knowledge area(s)**

Safety & Security  
ATM & Airport Simulation &  
Validation

### **Descriptor(s)**

SAFETY  
VALIDATION  
FEEDBACK





NLR-TP-2007-978

## Study of the Quality of Safety Assessment Methodology in Air Transport

M.H.C. Everdij and H.A.P. Blom

This report is based on a presentation held at the 25th International System Safety Conference (ISSC), 'Engineering a Safer World', Baltimore (Maryland), U.S.A., 13-17 August 2007.

The contents of this report may be cited on condition that full credit is given to NLR and the authors.

Customer                    NLR  
Contract number        ----  
Owner                      NLR  
Division NLR             Air Transport  
Distribution              Unlimited  
Classification of title    Unclassified  
Approved by:

Author	Reviewer	Managing department
<i>AE</i> 21/12/2007	Anonymous peer reviewers	<i>W. Rutke</i> 91.

## Study of the Quality of Safety Assessment Methodology in Air Transport

M.H.C. Everdij; NLR Air Transport Safety Institute, Amsterdam, The Netherlands  
H.A.P. Blom, PhD; NLR Air Transport Safety Institute, Amsterdam, The Netherlands

Keywords: safety assessment method, validation, feedback support, air transport

### Abstract

Air traffic is expected to double within the next 20 years, which requires large changes in airspace structure and organisation. The aim of this paper is to determine what is required for a safety validation approach to be appropriate for handling such large changes, and to test if existing approaches already satisfy this requirement. To this purpose, a consolidated set of indicators has been developed, each of which describes one required aspect of a safety validation approach for large changes, and the complete set describes all aspects. The consolidated set includes indicators related to the scoping of safety validation, to coverage of certain aspects of the operational concept, to risk assessment, to feedback to concept of operations development, to organisation of safety assessment, and to supporting decision and policy makers. Subsequently, the indicators are used to evaluate current safety assessment approaches in air transport. This shows needs for development of approaches towards safety assessment of large changes in air transport operations.

### Introduction

Air traffic is expected to double within the next 20 years, which requires large changes in airspace structure and organisation. Such large changes require an appropriate safety validation<sup>1</sup> before such changes can be implemented. Obviously, a safety validation can be done in different ways, and the quality of the result will depend on how the safety validation process is done, on the quality of the input and the experts used, which safety issues were evaluated, and which aspects of the operation were sufficiently covered. In order to take advantage of safety validation support methods from other safety critical industries, (ref. 8) developed a database of safety methods (ref. 10). This database contains over 600 methods from various domains, including aviation. Reference 8 also reports practical experience in using this data base for the search of relevant safety methods. Based on this kind of experience we felt that a complete and consistent search of such database asked for the development of an appropriate set of safety validation quality indicators.

Using reference 2 as a starting point, in reference 9 we studied principles for a safety validation framework for major changes in air transport operations. The systematic development of a set of safety validation quality indicators made part of this study. The aim of the current paper is to explain the process Phases that we followed for the development and application of this set of indicators. In Phase 1, potentially relevant candidate indicators are identified. In Phase 2, these candidate indicators are analysed. In Phase 3, these indicators are consolidated and further detailed. In Phase 4, the set of consolidated indicators is used to identify needs for improvement of an established approach. In Phase 5, the safety methods database is searched for methods to support these needs for improvement. Phases 1 through 4 are described in the current paper. Phase 5 has not yet been addressed, and will be subject of follow-up research.

### Phase 1: Identification of Candidate Indicators

The indicator development process started with the identification of as many as possible candidate indicators, using sources from literature and through conducting a brainstorm session with experts. The brainstorm session provided 91 candidate indicators. Participants of this brainstorm were qualified representatives of the Netherlands Ministry of Transport, of the Netherlands Air Traffic Control service provider, of the Netherlands National Supervisory Authority, and of Eurocontrol (Brussels and Bretigny-sur-Orge), complemented with NLR experts on safety validation and air transport operational concept development.

---

<sup>1</sup> Commonly, 'validation' is defined as answering the question "are we building the right system?", as opposed to 'verification', which is defined as answering the question "are we building the system right?"

The literature sources used are references 1, 2, 4, 5, 6, 14 and 15. This provided 108 candidate indicators. Reference 1 contains a directory of evaluated techniques to assess the dependability of critical computer systems. Reference 2 identifies the need to develop an ICAO global integrated system wide approach in aviation safety assessment that integrates all components of the aviation system in a balanced way, and lists indicators that this approach should satisfy. Reference 4 gives information quality guidelines to US Federal agencies, issued by the US Office of Management and Budget (OMB). Reference 5 contains a collection of hazard analysis and safety assessment techniques for use in the ATM/ATC domain. Reference 6 contains a collection of evaluated (technical) system safety analysis techniques. Reference 14 is a human reliability assessors guide, providing indicators for the evaluation of human reliability assessment techniques. Reference 15 contains a collection of evaluated techniques dealing with identifying human errors in high risk complex systems.

Finally, 16 additional candidate indicators were identified by the authors of this report during the course of the study. The result was a list of 215 candidate indicators, amongst which some doubles. In order to keep the consolidation process of this long list of candidate indicators manageable, the candidate indicators were divided over the five initial groups described in Table 1.

Table 1 — Initial Groups of Candidate Indicators

Group	Candidate indicator group description
Feedback	Related to interactions with, and effective feedback to operational concept design; also including communication of safety results with operational concept designers; related to appropriate coverage by the safety assessment framework of certain operation aspects (like human factors, interactions between operation elements, procedures, etc.), essential to make the feedback effective; and related to interactions with decision makers, and other authorities.
Application	Related to the kinds of changes to air transport operations that the safety assessment framework should be able to handle, e.g. major changes, Single European Sky, etc.
Acceptability	Related to transparency and international acceptability (and harmonisation), e.g. by ICAO.
Compliance	Related to compliance with certain safety assessment framework regulations; and related to making effective use of existing methods and norms.
Methods	Related to particular safety assessment steps that the safety assessment framework should include; and related to more general qualities that the safety assessment framework should have, e.g. be systematic.

### Phase 2: Analysis of Candidate Indicators

In the second Phase, for each of the five initial groups, the potentially relevant candidate indicators are summarised and analysed through an iterative process, in which review comments of the project Supervision Group<sup>2</sup> have been incorporated.

Analysis of Feedback Group: This subset covers candidate indicators that are related to the air transport operation design and lifecycle, and those related to interactions with decision makers and other authorities.

In summary, this group of candidate indicators express that a safety assessment framework should be able to provide effective safety results after each major air transport operation design stage, including the early stages. Here, the word “effective” is made more concrete by:

- The safety assessment should fit in the planning of operation design. In particular: Safety assessment results should be produced relatively quickly, in order not to slow down the operation design process; The framework should be able to produce results even if the input data is of low quality and /or is subject to change, e.g. as applicable in the earlier operation development stages, e.g. by inclusion of sensitivity analysis, and by allowing for regional flexibility;
- The framework should give insight into: Whether the air transport operation satisfies a safety design target, although in the first stages of operation lifecycle, the feedback could be in terms of ‘safety targets’. In particular, the framework should be quantitative since showing safety priorities is far more efficient with quantified models which in the end enable a cost-effectiveness evaluation; Where the design can be further developed while improving safety; A feeling for the

<sup>2</sup> The indicator development process was supported and reviewed by a Supervision Group with representatives of Eurocontrol, DGTL, LVNL, and Inspectorate V&W.

- impact of major changes on safety; The overall contributions to aviation risk; The relative importance of different accident categories and the causal factors to risk; The contributions in both causing and preventing aviation accidents; The relative importance of the different phases of the gate-to-gate cycle, e.g. the effects of strategic versus tactical measures; How much room there is for meeting the safety targets; The effects of interdependencies between different sub-systems.
- The framework should be able to produce results that are easily communicated to the air transport operation designers; The results should be transparent to, be acceptable to and connect to the intuition of the operation concept designers (“perceived validity”).
  - The safety assessment should also: Investigate changes in the interactions between different system elements; Investigate changes in the interfaces between different system elements; Investigate procedures, dynamical aspects, training issues, humans, equipment; (hardware and software), organisation, managerial aspects; Support a total systems approach and consider all specific phases of flight; Incorporate communication and navigation performance; Cover strategic, pre-tactical and tactical control elements; Address the identification of non-functional hazards; Cover Safety Management and monitoring activities.

Discussion and synthesis: Key to the safety assessment framework needed is that it should provide effective feedback to operational concept development, during all operation lifecycle stages, with special emphasis to the earlier stages. A second important aspect is related to timing. The safety assessment should fit in the planning of the design, and must therefore not need too much time to produce results. A third aspect, related to the second, is that the framework should be able to produce effective results even if the input is subject to change. These three aspects can be translated into concrete indicators, **Feedback and communication**, **Flexibility**, and **Information / data needed**. A fourth aspect is related to transparency of the result to be provided to the operation designers. For this, we include an indicator on how well the safety assessment process ensures that the result is transparent. Safety assessment results that are conflicting with the intuition of experienced domain experts may be acceptable if the safety assessors can convincingly explain why. For major changes, the safety effect may not at all be predictable, even by experienced experts. Hence, feedback and communication with domain experts are important qualities for the safety assessment framework, and this is translated into indicator **Transparency of results**. A fifth aspect is the depth of the assessment and the depth of coverage of all concept elements by the framework. An important element will be that the framework should cover all interactions between the concept elements. Therefore, this is formulated as a separate indicator, **Interactions and environment**. A sixth aspect concerns breaking down a safety target to the level of detail required, indicator **Safety Target breakdown**. Three additional aspects are related to interactions beyond operation designers, i.e. those related to decision makers, to regulatory authorities, and to safety oversight. These are translated into three more indicators, **Support to decision makers**, **Support to regulatory authorities**, and **Support to safety oversight**.

Analysis of Application Group: This subset contains candidate indicators which are related to kinds of changes to air transport operations. In summary, these candidate indicators make clear that the framework should be flexible, and applicable to a wide range of (future) air traffic and flight operations; But in particular: It should be applicable to the assessment of major changes in air transport operation; It should be applicable to the Netherlands airspace situation (both civil and military), or at least in a wide context like an extended airspace (i.e. with specific airports, including all approach and departure procedures, air traffic routes, separation criteria, and system performance); Ultimately, the framework should be applicable at an international level (e.g. Single European Sky); And be able to handle National implementation of operational concept designs that have been developed at International level.

Discussion and synthesis: Key indicator is that the framework should be applicable to the safety assessment of major changes in air transport operation, **Transparency regarding applicability**. Typically, major changes will involve replacement or change of procedures or re-organisation of air traffic control and/or airspace. Therefore, the framework should be able to cover technical systems (hardware and software), human factors, procedures, organisation (including culture), and institutional elements, implicitly including common causes. Here, regarding human factors, there are two distinct issues to consider: Human factors from the human operator perspective, and human factors from the perspective of the safety risk of conducting the operation considered. The latter includes human error. This can be translated into the following set of consolidated indicators: **Coverage of technical systems**, **Coverage of human factors for risk**, **Coverage of human factors for human**, **Coverage of procedures**, **Coverage of organisation**, and **Coverage of institutional elements**.

Analysis of Acceptability Group: This subset of potentially relevant candidate indicators are related to transparency and international acceptability. In summary, these candidate indicators support the following:

- The framework should be able to find support nationally and internationally. More concrete indicators are: Communicability to outside world; Accessibility to international forums; Acceptability by Inspectorate V&W and DGTL as a means to develop safety cases; Acceptability for international standardisation; Acceptability to regulatory and political bodies; Acceptability to scientific community; Expert review;
- The framework should be acceptable to assessors. According to one candidate indicator, this is most likely met by methods that require least resources and which have been most extensively applied. Hence, related indicators are: Resources required; Mastery required; User-friendliness and triability; Availability of supporting tools; Maturity;
- The framework should be transparent to experts. This is made more concrete by: The safety assessment results should be traceable; here some indicators related to traceability and verifiability are: Auditability; Documentability; Observability; Availability; Comparative validity; Consistency; Structuredness; Reproducibility; Transparency; Compatibility; Quality, i.e. Utility, Objectivity and Integrity; Perceived validity.

Discussion and synthesis: A key indicator is that the framework can collect support for the approach, nationally and internationally. To do so, not only technical but also political aspects need to be addressed. For example, several organisations have already invested in a safety assessment framework of their own, and will want to see that one implemented internationally, rather than another one. On the other hand, if the new framework can really show to have advantages above existing ones, the support will be found easier. One indicator that may support this asks whether the method is able to withstand criticism, **Criticism**. Another set of candidate indicators is related to acceptability to safety assessors, and the people who are going to pay for performing the safety assessment of a new operation. They will be interested to know what applying the framework requires in terms of resources, e.g. number of experts required, including their training. This becomes indicator **Resource requirements (equipment and personnel)**. The third set of candidate indicators is related to transparency. The problem is that transparency in itself may be hard to measure. It is strongly dependent on the expertise and experience of the person reviewing the method and results. More concrete candidate indicators related to transparency are listed in the summary above. It may not be considered logical or required to include all of these indicators in our eventual list. A relevant selection may be sufficient. Two indicators that may be considered most relevant are Documentability, **Documentability of process steps**, i.e. the degree to which the technique lends itself to auditable documentation, and Consistency, Indicator **Consistency**, which measures that if the method is used on two occasions by independent experts, reasonably similar results are derived. Documentability may also cover auditability and observability to some extent, and Consistency may also cover structuredness and reproducibility to some extent, such that these four additional indicators are also more or less covered.

Analysis of Compliance Group: This subset of potentially relevant candidate indicators are related to international norms and developments. In summary, the candidate indicators support the following:

- Framework should connect with International Norms. To be more concrete, Framework should be compliant to Eurocontrol's ESARR 4 and the EU's Common Requirements; Framework should be compatible to existing regulations; Framework should be adaptable to future regulations.
- Framework should make effective use of existing national and international methodological developments. Some concrete example indicators are: Ease of integration with / compatibility with other approaches; Development potential.
- However, the development of the framework should also be open to innovation and not be biased towards existing methods: it may be that currently existing methods are not satisfactory;
- Framework should give a large 'push' to international developments, for example Eurocontrol risk classification scheme development.

Discussion and synthesis: An important indicator, which is also necessary for international acceptability, is that the framework should be compliant with international norms and regulations. There are several such regulations, and in addition, some are under development. Some international regulations are ESARR 4, the Common Requirements (CR) of the EU, and regulations posed by ICAO. There are relevant points of criticism regarding ESARR 4 and the Common Requirements, and it is possible that the regulations will be updated in the near future to take this criticism into account.



However, throughout the states, ESARR 4 is regarded as a standard, and in many places, ESARR 4 and / or CR compliance is considered essential for acceptability. So, we definitely should take them well into account. In addition, there may be other requirements, e.g. aircraft-related certification / performance requirements that may be relevant for air traffic. This can be translated into one indicator **Compliance to ESARRs, CR, ICAO**. The indicators related to making use of and giving a large push to existing developments will not be translated into a specific indicator here, but will be incorporated more indirectly in the framework to be developed: This report will take some internationally well regarded methodologies as reference points for the framework to be developed. In addition, we intend to look beyond these existing methodologies and try to include some innovative ideas to complement them towards satisfying the project objectives. This, and the efforts to find international and national support for the framework, should ensure a large push to international developments.

Analysis of Methods Group: This subset of potentially relevant candidate indicators are related to safety assessment methods used by the framework. In summary these candidate indicators support the following:

- The framework should cover the seven (or eight) stages of a generic safety assessment process (ref. 11 or ref 6): 1. Scoping the assessment; 2. Learning the nominal system; 3. Identifying hazards; 4. Combining hazards; 5. Evaluating risk; 6. Supporting risk mitigation; 7. Monitoring / verifying actual risk; 8. Feedback and communication (has to be part of each of the other stages).
- With some particular remarks to these steps being: The safety target should be set outside the safety assessment; The framework should support a good breakdown of the risk classification scheme into smaller risks and support a definition of safety targets compatible with the scope and level of detail in which the concept has been developed; Framework should include verification and validation steps.
- Other qualities: Experts required; Degree of decomposition; Thoroughness; Systematism; Progressiveness; Application of the framework should lead to improvements in the framework (and in the way the framework structure is organised); Level of safety expertise required.

Discussion and synthesis: The generic safety assessment process referred to above consists of seven stages, with two feedback loops: One feedback concerns iterations to previous stages, and the other concerns feedback to operations, assessment and design. This latter stage may also be referred to as an “eighth” stage, although it should be part of each of the other stages. It is important that each of these stages gets proper treatment within the safety assessment framework, since each will have their own effects on operation concept design. Therefore, we decided to translate each of the first seven stages as a separate indicator, **Scoping the assessment, Learning the nominal operation, Identifying hazards, Combining hazards, Evaluating risk, Supporting risk mitigation, Monitoring / verifying actual risk**. The eighth stage is also very important, but it should be covered by the Group Feedback indicators. In addition, the stage on evaluating risk is covered by two separate indicators: one to evaluate the risk according to the identified scenarios, and a separate one to cover nominal risk, i.e. risks during normal operations within a hazard-free scenario, **Coverage of nominal risk**. The summary above also lists some remarks to particular stages. The first is on the safety targets. These targets should be set outside the safety assessment. The second remark is on a breakdown of the overall safety target into risk budgets for suboperations. Identification of safety targets and their breakdown is part of the scoping stage of the safety assessment process, but to stress its importance, it is covered by a separate indicator at Group Feedback (Safety Target breakdown). The third remark is on verification and validation. These should be part of each major stage in the safety assessment process, hence they could be covered by all stages. However, they can also be covered by indicators from a slightly different perspective: During the various stages of safety assessment, approximations (or assumptions) need to be made. This already starts during hazard identification: there is an implicit assumption that all important hazards are identified. But also during the risk modelling and risk evaluation, several approximations need to be adopted. Verification and validation steps should check if these are reasonable assumptions and if the deviation from reality is not too large. In order to capture this in our framework evaluation, we propose one additional indicator, namely, the extent to which the framework identifies and evaluates the approximations made, **Approximations analysed**, and an additional indicator on transparency: **Transparency of safety assessment process**. Finally, the summary above lists some additional qualities that could be desirable for the framework. Some of these qualities have already been partly covered by the indicators related to Group Feedback and Acceptability. Only the last one may be relevant: Level of safety expertise required by the framework, **Level of safety expertise required**.





### Phase 3: Consolidation of Indicators

Through the Phase 2 analysis, 32 indicators have been developed. In Phase 3, these are further described and detailed, put in a more logical order, divided into six classes, and finally numbered CI-01 through CI-32 (where CI denotes consolidated indicator). See Tables 2 through 7 below. The last column indicates for traceability reasons the initial group in which the indicator was developed. The six classes are developed by an iterative process, and are based as much as possible on different stages in safety assessment of a safety critical operation (ref. 11).

Table 2 — Indicators Related to the Scoping of Safety Validation

CI-01	Information / data needed	How well can the method produce effective results if there is only limited input information available from operational concept designers?	Feedback
CI-02	Scoping the assessment	How well does the framework handle Scoping the assessment stage, which entails writing a safety plan that specifies the scope of the safety assessment and outlines a “route map” for the safety assessment? How well is the safety target defined outside the safety assessment?	Method
CI-03	Safety Target breakdown	Does the method support a breakdown of the safety target to the level of detail required, during all stages of the lifecycle?	Feedback
CI-04	Learning the nominal operation	How well is learning of the nominal operation handled, i.e. learning to understand the operation and systems as they should work or function?	Method

Table 3 — Indicators Related to Coverage of Certain Aspects of the Operational Concept

CI-05	Identifying hazards	How well are hazards identified? How well does the method support the identification of future hazards, i.e. hazards that may not be known yet, but may occur in future operations? Does the hazard identification cover all aspects of the future operation?	Method
CI-06	Coverage of technical systems	How well are technical systems (hardware and software) covered, including technical systems that can be expected for future operations?	Application
CI-07	Coverage of human factors for risk	How well are human factors covered from risk perspective, including human factors that can be expected for future operations?	Application
CI-08	Coverage of human factors for human	How well are human factors covered from human perspective, including human factors that can be expected for future operations?	Application
CI-09	Interactions and environment	How well is the coverage by the method of interactions between multiple agents in the operation (e.g. air traffic controller, pilot, military ATM, navigation and surveillance equipment, search and rescue), with the environment of the operation?	Feedback
CI-10	Coverage of procedures	How well are procedures covered, including procedures that can be expected for future operations?	Application
CI-11	Coverage of organisation	How well is the organisation within and between stakeholders covered, including organisation that can be expected for future operations?	Application
CI-12	Coverage of institutional elements	How well are institutional elements covered, including institutional elements that can be expected for future operations?	Application

Table 4 — Indicators Related to Risk Assessment

CI-13	Combining hazards	How well are hazards combined, connected to safety-related scenarios and evaluated?	Method
CI-14	Evaluating risk	How well does the framework evaluate the risk according to the identified scenarios?	Method
CI-15	Coverage of nominal risk	How well does the method address the risks during normal (nominal) operations, i.e. the systems and procedures are	Method

		designed and a hazard-free scenario is being considered?	
CI-16	Approximations analysed	How well does the framework identify and evaluate approximations made with respect to reality?	Method

Table 5 — Indicators Related to Feedback to ConOps Development

CI-17	Feedback and communication	How well is feedback (if any) communicated with operation design?	Feedback
CI-18	Supporting risk mitigation	How well does the framework support the identification of effective risk mitigation strategies?	Method
CI-19	Monitoring / verifying actual risk	How well does the framework support the monitoring and verification of actual risk?	Method

Table 6 — Indicators Related to Organisation of Safety Assessment

CI-20	Resource requirements (equipment and personnel)	Is the level of resources needed reasonable for the results delivered (where resources refers to number of personnel, their training, availability and length of their time required by the study, as well as equipment and administrative support requirements)?	Acceptability
CI-21	Criticism	Is the method able to withstand criticism?	Acceptability
CI-22	Level of safety expertise required	How well does the method pose requirements on the designated safety assessor to have the proper operational safety expertise background?	Method
CI-23	Documentability of process steps	What is the degree to which the framework lends itself to auditable documentation?	Acceptability
CI-24	Consistency	How well is the consistency of the use of the framework, such that if used on two occasions by independent experts, reasonably similar results are derived?	Acceptability
CI-25	Compliance to ESARRs, CR, ICAO	How well is the level of compliance to: ESARRs, EC's Common Requirements, ICAO requirements, or other international requirements (e.g. aircraft-related certification / performance requirements relevant for air traffic)	Compliance
CI-26	Flexibility	In case of a modification in the operational concept description when the safety assessment is already ongoing: How much additional time / effort is required to update the safety assessment accordingly?	Feedback

Table 7 — Indicators Related to Supporting Decision and Policy Makers

CI-27	Transparency regarding applicability	To what extent does it become clear which applications (e.g. air transport operations, aircraft flight, runway incursions, Single European Sky) are accommodated?	Application
CI-28	Transparency of results	How well is the transparency of the results, where transparency is defined as Understandable, Traceable, and Well documented	Feedback
CI-29	Transparency of safety assessment process	To what extent are the steps in the safety assessment process or framework transparent to the safety assessor?	Method
CI-30	Support to decision makers	How well does the method provide support to decision makers?	Feedback
CI-31	Support to regulatory authorities	How well does the method support presentation to and communication with regulatory authorities?	Feedback
CI-32	Support to safety oversight	How well does the method support checking / verification by safety oversight?	Feedback

#### Phase 4: Needs for improvement in air transport safety validation

A logical next step after the consolidated indicator development process would be to go start developing a safety assessment framework that satisfies all these CIs. However, the development of a new safety assessment framework from scratch is challenging, and could take many years. A more

realistic approach is to first evaluate if there are established safety assessment methods that already satisfy the CIs. Since the method we are looking for is aimed at getting international acceptability, a most logical choice for this is to start with a few safety assessment methods that are already widely carried and widely used within the international aviation community.

There are several advantages of using established methods as reference point: First of all, there is no reason to try to invent the wheel again: Established methods typically are the result of many years of thinking and developing, and may already satisfy several of the CIs. In addition, established methods have been applied by many, and to many practical situations, which gives us practical insight in the advantages and disadvantages of the method in relation with their interactions with operational concept design. And finally, established methods usually already have support from part of the international community, hence this part does not have to be convinced anymore of using the method as a reference.

This section aims to identify to what extent the CIs developed above can be satisfied by using a safety assessment method that is established within Eurocontrol, and to identify needs for improvement in order for it to be used for the safety validation of major changes in air transport operations. As a fortunate side effect, this exercise should also provide insight into the effectiveness, usefulness and completeness of the developed CIs. This evaluation process has been done twice: the results of the first cycle have been used to improve the CIs. The results of the second cycle are presented next.

Evaluation of Established Safety Assessment Approach: The first step was to identify one or more established safety methods, by using the criterion “widely carried and widely used within the international aviation community”. There appeared to be several candidates, but for this paper we present the analysis results of only one, i.e. the safety validation method as applied to a final approach operation that is supported by GBAS (Ground-Based Augmentation System); the results of this safety validation are documented in references 12 and 13.

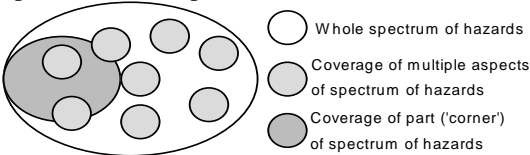
The safety validation method used is based on EATMP SAM (ref. 3), which is a safety assessment methodology widely used by European ANSPs (ref. 6). The reason why we opted to evaluate references 12 and 13 rather than reference 3 itself is explained by the formal methodology description leaving freedom in how to execute the different methodology steps. There is a lot of guidelines material, and in addition, there is course material, and safety case material for several actual applications. This yields that both in theory and in practice still different choices can be and are being made (e.g. there is freedom to choose a particular technique for an analysis step, in favour of another), and these choices affect the evaluation against the consolidated indicators. For this reason, the approach taken in this step is to select one representative application (a practice) of the method to a particular operation, and to only evaluate the method as conducted by this particular application. Advantages of selecting a practice rather than theory are that only practice can show if the theory works, and that the number of possible choices for conducting the method has been reduced to only one. A disadvantage is that the choices made for applying the method to different practices may not always be the same. However, after weighing all advantages and disadvantages of this and other approaches, it was identified as the option to follow.

We also note that although the authors of references 12 and 13 consider this example representative for reference 3, we should note that there are larger changes in air transport operations needed than GBAS alone, hence an evaluation of references 12 and 13 against CIs that were developed for major changes involves the need for expert extrapolation and judgement. In such cases the outcomes have been discussed with Eurocontrol experts in order to avoid a significant bias by the evaluators.

Another issue to be covered is what scoring system to use to present the evaluation results? In the beginning, we tried out several kinds of scoring mechanisms expressing the extent to which the method satisfied the indicator. For example, + / 0 / – was expressing that the method scored well / average / not sufficiently well. The drawback of this approach is that it distracts the evaluation away from the objective of identifying areas for improvement. The alternative scoring of \*\*\* / \*\* / \*, indicating level of coverage by the method of an indicator has the same drawback. A third version was to identify suggestions for improvement of the method to better cover the indicator. These suggestions needed to be concrete and a positive approach needed to be maintained. Because the third option was seen as the most constructive, it was selected as the option to follow.

**Needs for Improvement Initially Identified:** Next, the safety method used in references 12 and 13 was analysed and evaluated against each of the CIs (except those in Table 7), and in case the method did not completely satisfy a CI, a need for improvement is initially identified. It appeared that often, one need for improvement could be related to more than one CI. Therefore, we subsequently consolidated the set of needs for improvement and related each need to a set of CIs that would further profit if this need would be satisfied. The table below provides the needs for improvement identified, together with the subset of CIs that would profit. The last column of the table gives the possible consequences if this need is not satisfied, which motivates why the need for improvement is very relevant to provide a sound safety case for the major changes in air transport operations envisioned for this paper.

Table 8 — Needs for Improvement of Safety Validation According to GBAS to Further Meeting the Developed CIs

Need for improvement	Related indicators	Consequences if need is not satisfied
Organise training curricula to deliver high-level safety experts, in which safety experts are taught to have a responsibility towards the safety methods they are using	CI-20, CI-22	If this is not done, consequence may be that methods are used without attention to their drawbacks. Another consequence may be that safety experts go too much into technical detail, and as a result do not stimulate the operational concept designers to think beyond technical detail either.
Work with higher-level operational concept description that is described in a goal-oriented way per human agent, and that covers organisational and institutional aspects	CI-02, CI-03, CI-04, CI-06, CI-10, CI-11, CI-12, CI-27	If this need is not satisfied, consequence is that organisational and institutional aspects will not be considered, and hazards related to these aspects are not taken into account. Another consequence may be that there is no harmonisation between safety target setting/scoping and the joint goals setting of the stakeholders involved in the operation.
Use hazard identification methods that complement the ones used, in order to push the boundary between ‘imaginable’ and ‘unimaginable’ hazards	CI-01, CI-05, CI-07, CI-09, CI-10, CI-11	<p>If this need is not satisfied, consequence is that there is no view on multiple aspects of the whole spectrum of hazards, only hazards in the ‘corners’ (see figure) may be considered. Areas of the spectrum of hazards that are not identified will not be taken into account in the operational concept.</p> 
Avoid the need for fixed event sequences in scenarios and cover hazards more explicitly; go beyond human error thinking and address human factors from a human perspective, cover interactions between multiple agents of the operation, and with the environment, give appropriate attention to goal-directed use and prioritisation of procedures and bring organisational and institutional issues into account	CI-06, CI-07, CI-08, CI-09, CI-10, CI-11, CI-13, CI-14, CI-15, CI-26	If this need is not satisfied, consequence is that there is no way to analyse particular types of hazards, specifically hazards that go beyond technical failures and human error. Without proper support to analyse these types of hazards, it is not possible to learn understanding their safety impact well, and to explain this well to operational concept designers. Hence these effects are not likely taken into account in the design.
Improve systematic safety data collection and safety performance monitoring	CI-16, CI-19	There are multiple consequences if this need is not satisfied: There will be less insight in current risks, hence less insight in benchmark for the future. Another consequence is that an opportunity is missed to identify several interesting hazards that occur today, but that have not been identified yet. These hazards, if not identified, can also not be monitored for future designs.
Identify all assumptions adopted during the risk assessment,	CI-16, CI-17, CI-28, CI-23	If this need is not satisfied, consequence is that important safety effects that are hidden in assumptions

Need for improvement	Related indicators	Consequences if need is not satisfied
implicitly and explicitly, and evaluate them against reality; identify what has not been covered by the safety assessment and communicate this to the operation concept designers		remain invisible. In addition, a missing proper analysis and explanation of all assumptions may lead to miscommunication with operational concept designers, who may be led to believe that important design issues are 'reasoned away' in assumptions.
Leave the identification of operational mitigating measures to operation designers, but provide them with proper support	CI-17, CI-18	First of all, safety experts may have a less realistic insight in the operational consequences (i.e. from a pilot or controller point of view) of certain mitigating measures than operation designers do. In addition, safety experts who have identified mitigating measures themselves, may find it difficult to do an objective safety analysis of these same measures.
Study compliance to CR and ICAO requirements	CI-25	Consequence of non-compliance to these requirements may be that there is less harmonisation and standardisation of advanced operational concept development, e.g. between countries. In addition, it should be taken into account that the requirements themselves may also need to evolve through time; e.g. current safety design targets may not be applicable to future operations.
Study how one can evaluate a method against the following indicators: Criticism, Consistency, Transparency, Support to decision makers, Support to regulatory authorities, Support to safety oversight	CI-21, CI-24, CI-29, CI-30, CI-31, CI-32	If these indicators are not well understood, there is a chance that safety methods do not satisfy these and other indicators. Consequence may be that e.g. important key decision makers will get insufficient information to review safety cases, and to make proper decisions.

### Concluding Remarks

This paper developed a set of quality indicators for safety validation of major changes in air transport operations. Through a process of analysis, evaluation, synthesis and review, a consolidated set of 32 indicators was developed for measuring to what extent a given safety validation method can be used to develop a good safety case for such major change. These indicators were initially verified and tested on an established safety assessment approach. As a result of this, we identified in an objective way several needs for improvement of the established safety assessment approach in support of the safety validation of future major changes in air transport. This forms a clear illustration of the practical use of the developed set of consolidated indicators of safety validation quality.

### References

1. P.G. Bishop (editor), Dependability of critical computer systems - Part 3: Techniques Directory; Guidelines produced by the European Workshop on Industrial Computer Systems Technical Committee 7 (EWICS TC7), Elsevier, 1990
2. DGCA, Proposed paper on the development of a global gate-to-gate safety assessment methodology, Paper presented by the Netherlands at 121st ICAO meeting of Directors General of Civil Aviation, DGCA/121-DP/17, 15 June 2004.
3. EATMP SAM, Air Navigation System Safety Assessment Methodology, SAF.ET1.ST03.1000-MAN-01, Edition 2.0, 30 April 2004.
4. EDA (Economic Development Administration) Information Quality Guidelines, Part I: Background, Mission, Definitions, and Scope, [http://www.eda.gov/ImageCache/EDAPublic/documents/pdfdocs/infoqualguidecas\\_2d9\\_2d26\\_2epdf/v1/infoqualguidecas\\_2d9\\_2d26.pdf](http://www.eda.gov/ImageCache/EDAPublic/documents/pdfdocs/infoqualguidecas_2d9_2d26_2epdf/v1/infoqualguidecas_2d9_2d26.pdf)
5. M.H.C. Everdij, M.B. Klompstra, H.A.P. Blom, O.N. Fota, MUFTIS work package report 3.2, final report on safety model, Part I: Evaluation of hazard analysis techniques for application to en-route ATM, Report NLR TR 96196 L, 1996

6. R.A. Stephens, W. Talso, System Safety Analysis handbook: A Source Book for Safety Practitioners, System Safety Society, 1st edition in 1993, 2nd edition in 1997.
7. M.H.C. Everdij and H.A.P. Blom, Safety Assessment Methodologies, CAATS Deliverable D1.4 Safety Report, <http://www.caats.isdefe.es/>, 2006.
8. M.H.C. Everdij, H.A.P. Blom, B. Kirwan, Development of a structured database of safety methods, Proc. 8th Int. Conf. on Probabilistic Safety Assessment and Management (PSAM8), May 2006, New Orleans, USA.
9. M.H.C. Everdij, H.A.P. Blom, J.W. Nollet, and M.A. Kraan, Need for novel approach in aviation safety validation, Second Eurocontrol Safety R&D Seminar, October 2006, Barcelona, Spain.
10. M.H.C. Everdij and H.A.P. Blom (editors), Database of over 600 safety assessment methods and techniques from various industries, Maintained by NLR, <http://www.nlr.nl/documents/flyers/SATdb.pdf>.
11. FAA/EUROCONTROL, ATM Safety Techniques and Toolbox, Safety Action Plan-15, Issue 1.0, 2005, [http://www.eurocontrol.int/eec/public/standard\\_page/safety\\_doc\\_techniques\\_and\\_toolbox.html](http://www.eurocontrol.int/eec/public/standard_page/safety_doc_techniques_and_toolbox.html)
12. GBAS FHA, Category-I (CAT-I) Ground-Based Augmentation System (GBAS) Post Concept Functional Hazard Assessment, Edition V1.1, May 2004, Proposed issue, Classif. Restricted, 2004.
13. GBAS PSSA, Category-I (CAT-I) Ground-Based Augmentation System (GBAS) Post Concept PSSA Report, Edition V1.0, 18 Nov 2005, Proposed issue, Classif. EATM, 2005.
14. P. Humphreys, Human reliability assessors guide, Safety and Reliability Directorate UKAEA (SRD) Report No TRS 88/95Q, October 1988.
15. B. Kirwan, Human error identification techniques for risk assessment of high risk systems – Part 1: Review and evaluation of techniques, Applied Ergonomics, Vol 29, No 3, pp. 157-177, 1998.