SO IT'S RELIABLE BUT IS IT SAFE?
A MORE BALANCED APPROACH TO ATM SAFETY ASSESSMENT

Derek Fowler, Gilles Le Galo, Eric Perrin, EUROCONTROL, Brussels, Belgium
Stephen Thomas, Entity Systems Ltd, UK

Abstract

There is a view, reflected in some safety-critical software research in the USA and in experience of safety assessments in Air Traffic Management in Europe, that safety stems largely from reliability of systems.

Whilst that view may be appropriate for systems that are simply inherently dangerous (nuclear power plants for example), it would be far too narrow for the more general case of systems that have a specific safety-related purpose to fulfil – for example, car airbags, flight control systems and ATM – since it would exclude consideration of the positive contribution that such systems are required to make to the safety of their operational environment, host system etc.

The paper discusses the implications of this issue for ATM and, whilst acknowledging that this limitation may not have been a major problem in the past (because of the hitherto gradual evolution of ATM systems) it explains why a pre-occupation with system failure (at the expense of functionality and performance) cannot be sustained in the face of more radical changes being considered for ATM over the next 20 years or so.

The paper then presents a new framework, for a broader approach to ATM safety assessment, in the form of a high-level safety argument. It covers what are known as success and failure approaches: the former is concerned with ensuring that ATM systems / services actually deliver the necessary reduction in aviation risk (ie contribute to the prevention of aircraft accidents) when functioning as required; whilst the latter is concerned with ensuring that there would be no significant increase in risk (ie cause, or fail to prevent, accidents) in the event of system malfunction.

The description of the framework is illustrated with examples from recent and current EUROCONTROL work on ATM operational concepts.

Introduction

Background

In a paper [1] on her review of major software-related accidents, for the US DoD, Nancy Leveson presents compelling evidence that software reliability had never been the cause of such disasters - on the contrary, the software had in every case performed in exactly the manner that it was designed to. The problem was that the software was designed to do the wrong thing for the circumstances under which it "failed". In identifying a need for a broader view of safety (ie beyond mere reliability), Professor Leveson poses the question why, in most software safety standards, so much emphasis is placed on processes to improve reliability whilst not ensuring that the resulting systems actually perform the intended function – ie allowing them to be "reliably unsafe".

For an everyday illustration of problem, take the case of a car airbag. Clearly, we would want it to be reliable - ie to operate when it is needed and not to operate when it is not needed. However, we would also want it to be effective (in preventing death / serious injury) when it does operate; this would depend on such things as its size, shape, construction and speed of deployment – ie on its physical and functional properties.

Overall, therefore, the (safety) case for fitting airbags to cars depends on their saving far more lives / preventing serious injury, when operating as intended, than any deaths / serious injury that they might cause in the unlikely event of spurious operation. If the balance between the risk-reduction benefits of successful operation and any disbenefits of induced risks arising from failure were not (significantly) in favour of the former, then the airbag could not properly be considered to be safe.

Situation in ATM

In the authors' wide experience of safety assessments, the problem identified in the context of safety-related software in [1] is prevalent in a general sense in European ATM, and it is quite common for safety assessments of ATM systems to address only how unsafe a system might be when it failed without assessing how safe it would be when it was working.

This may in part be due to a misunderstanding of EUROCONTROL safety regulatory requirement ESARR 4 [3], or at least to too narrow an interpretation of the term "ATM direct contribution to an accident" used therein. Since the primary purpose of ATM is to prevent accidents, it is the functionality and performance of an ATM system (people, procedures and equipment), as much as its integrity, which determines how safe, overall, the ATM service is. For example, from the results of the EUR RVSM Height Monitoring Programme [4], it can be deduced that the provision of vertical separation prevents about $2x10^6$ times as many accidents than loss of vertical separation causes[1]. Assuring the continuing effectiveness of vertical Separation Provision (in the absence of failure) must therefore be (and is!) one of the main goals of the on-going EUR RVSM Programme – see below.

In the past, a pre-occupation with failure (other than on RVSM and some other exceptions[2]), appears not to have presented much of a problem since the risk-reducing properties of ATM (ie its functionality and performance) had implicitly been established in ATM operational and technical requirements long before formal safety assessments were thought of, and until recently had changed very little.

However, it follows that as the underlying concepts in ATM become subject to change, the long-established assumption that ATM will implicitly continue to be as effective in reducing aviation risk cannot be sustained. This is borne out by experience on a number of current EUROCONTROL ATM development studies, and is also noted in a JAA-sponsored study of the safety of a future ATM Concept of Operations [2]. These have identified the need for a new (or at least much broader) safety-assessment approach, covering success as well as failure – ie functionality and performance as well as integrity[3] - in order to ensure the safety of such developments.

From a safety methodology point of view, the idea is not entirely new; the EUROCONTROL ANS Safety Assessment Methodology (SAM) [3] already identifies the need for success and failure approaches to safety assessment. However, the SAM currently does not provide much guidance on what the former entails and, as already said, the success approach is rarely applied.

The purpose of this paper, therefore, is to present a framework which addresses the need for a more balanced (success versus failure) approach to safety assessment, aimed at answering two key questions concerning the safety of a system:

- is it safe when operating to specification?
- will it still be safe if / when it fails in some way?

It starts by considering these two aspects of safety, from a theoretical perspective.

Theoretical Perspective

Basic Concepts

A good starting point for understanding the success and failure approach to safety assessment is IEC 61508 [6] - a widely recognised international standard on functional safety and designed to be generally applied (or specifically adapted where appropriate) for a wide range of sectors, including rail, aerospace, automotive and ATM.

IEC 61508 defines (implicitly, at least) a safety-related system (SRS), not as something that merely poses a threat, but rather as something that makes a net positive contribution to the safety of its environment (or host system).[4] *Safety Functions* are an abstract functional representation of an SRS and it is the realisation of these that make the required contribution to safety. *Necessary Risk Reduction* (NRR) is used to describe the amount of contribution required of the Safety Functions.

This is illustrated in Figure 1. The unmitigated risk (R_U) is that risk which exists in the environment (or host system), prior to the introduction of the SRS.

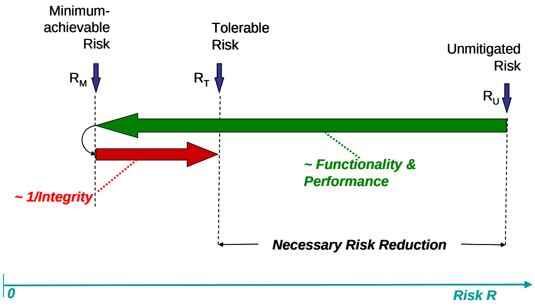


Figure 1 – Risk Reduction

The minimum NRR to be provided by the Safety Functions is, therefore, simply the difference between the value R_U and whatever is considered to be the maximum tolerable risk ($R_{T)}$ for the particular application.

[R_U-R_M] is the maximum, theoretical risk reduction that can be provided by the Safety Functions, assuming (hypothetically) that the physical SRS is entirely failure free. [R_U-R_M] is

[1] A simple comparison of the predicted accident rate and the rate at which aircraft have been recorded as being in "horizontal overlap" – see [4].

[2] Including those discussed later herein, and ADS-B (the subject of a separate paper [10])

[3] Hereinafter, the *integrity* is used instead of *reliability* as a general term covering all failure-related properties.

[4] Thus a nuclear reactor would not be thought of as being an SRS but its protection systems would be. It is shown later in the paper that ATM most certainly falls within the IEC 61508 definition of an SRS

determined entirely by the functionality and performance of the Safety Functions implemented in the SRS, in relation to the hazards and risks associated with its environment (or host system). It should be noted that R_M is not zero, because it is assumed that, due to the finite limits to its capabilities, there will always be some risks (however small) that the SRS is not able to eliminate , even when it is operating to its full specified functionality and performance.

Clearly, for a safe situation to be possible, R_M must be less than R_T. That being so, $[R_T - R_M]$ is the (risk) margin that allows for occasional failure of the SRS, and which determines ultimately the integrity[5] required of it.

Thus there is a direct relationship between an SRS's functionality and performance (since it is these properties that determine R_M) and its required integrity (determined by R_T and R_M)[6].

When the primary SRS's functionality, performance and integrity are sufficient only to meet the prescribed maximum-tolerable level of risk (R_T), it may be possible to add a secondary standby SRS in order to reduce the risk further, to a fully acceptable level $R_{A,}$ as illustrated in Figure 2.

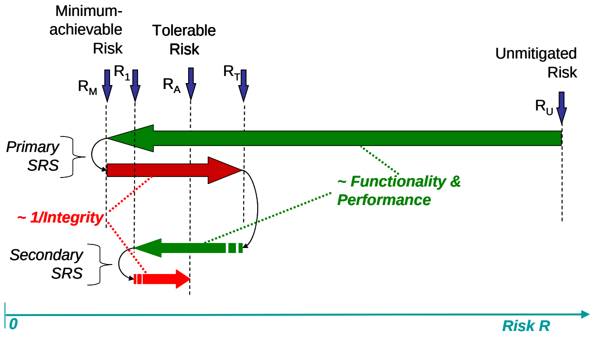


Figure 2 – Addition of Standby SRS

Derivation of Integrity

For SRSs that operate continuously, the required integrity is expressed as the maximum allowable failure rate (RF). It can be shown that this is given simply by:

$\boldsymbol{RF = R_T - R_M}$

For SRSs that operate in a *standby* role, however, the mathematics are somewhat different – this is because such SRSs are required to operate only when needed (ie on demand), and are often additional to other SRSs.

If, for example, the secondary SRS in Figure 2 operated in a standby role, then it can be shown that the allowable maximum probability of failure on demand (PF) is given by:

$\boldsymbol{PF = (R_A - R_1) / R_T}$

Implications

So what does all this mean in practice?

Firstly, the distinction (and relationship) between providing risk reduction (the *success* approach) and controlling risk increase (the *failure* approach) is vital, and safety is about establishing, and demonstrating, an appropriate relationship between the two[7] –.

Secondly, in order to assure the safety of SRSs, we need to specify both:

- Safety Requirements for functionality and performance, which provide the necessary risk-reduction potential of an SRS, and which are derived initially from hazards and associated risks inherent in the operational environment / host system – ie external to the SRS; and
- Safety Requirements for integrity, which limit the risk increase caused by failure internal to the SRS.

Having established the need for this broader, more balanced approach to safety assessment in general, the paper next discusses the relevance of this approach to ATM systems.

Relevance to ATM

A suitable starting point for explaining the application of success and failure approaches to ATM safety assessments is ICAO Doc 9854 [7]. This presents the ICAO vision of an integrated, harmonized and globally interoperable ATM system for the period up to 2025 and beyond. It includes a description of Conflict Management, a key component of the 'emerging and future' ATM Operational Concept, which:

- is aimed at **reducing**, to a tolerable level, the **risk** of collision between aircraft and other aircraft, fixed obstacles[8] etc; and
- is applied in three layers: Strategic Conflict Management, Separation Provision, and Collision Avoidance.

How this service-level concept works in practice, and relates to the underlying ATM system, can be seen from Figure 3 below.

[5] IEC 61508 uses the term Safety Integrity Levels (SILs) for this property.

[6] In the limit that R_M approaches R_T, so the integrity required of the SRS approaches infinity!

[7] A thesis that is entirely consistent with the Leveson view [1] that safety is not synonymous with integrity alone

[8] This places ATM firmly within the IEC 61508 definition of an SRS – see previous section

The input to this simple model is the air traffic, the existence of which represents hazards to, inter alia, one or more aircraft within it.

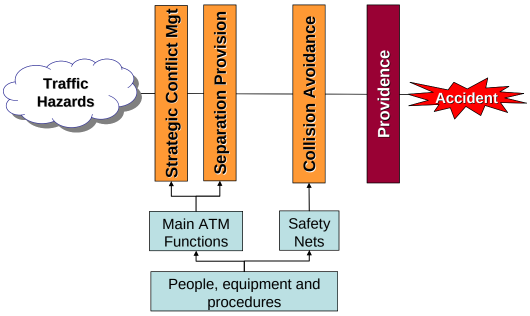


Figure 3 – Simplified Conflict Management Model

The three layers of Conflict Management identified in [7], can be thought of as *barriers* which prevent those hazards leading to an accident[9].

The barriers operate from left to right in a rough time sequence[10], and each barrier contributes to safety (ie reduces risk) by removing a percentage of the *conflicts*[11], which exist in the operational environment into which the ATM service is delivered, as follows.

The **Strategic Conflict Management** barrier is provided by the following main ATM functions:

- *Airspace design* which provides structuring of the airspace so as to keep aircraft apart spatially, in the lateral and/or vertical dimensions
- *Flow and Capacity Management* which mainly prevents overload of the Separation Provision barrier
- *Traffic Synchronisation* which involves the planning of the routing and timing of individual flights so that the aircraft, if they followed their planned trajectories, would pass each other without infringing the prescribed minimum separation.

The **Separation Provision** barrier is the second layer of Conflict Management and is the process of keeping aircraft away from each other, and from fixed obstacles etc, by at least the appropriate separation minima, by means of tactical intervention. Separation Provision may be the responsibility of an ANSP, the airspace user, or a combination of the two.

Separation Provision is based on *Tactical Conflict Detection* and *Tactical Conflict Resolution*, and is necessary due to the inherent limitations of Strategic Conflict Management in eliminating all conflicts.

Collision Avoidance is intended to recover the situation only when the previous two barriers have failed to remove conflicts to the point that applicable separation minima have been, or are about to be, infringed. It may be initiated by either:

- Controllers, often supported by ground-based safety nets such as STCA[12], MSAW and APW, or
- Flight Crew, often supported by airborne safety nets such as ACAS and GPWS.

Providence is the final barrier and simply represents the probability that, owing to the huge volume of airspace and the spread of traffic within it, aircraft involved in a given encounter, albeit in close proximity, would not actually collide. Although largely a matter of chance, Providence can be affected by such things as airspace design and traffic distribution, and its effectiveness generally decreases as the density of traffic increases with, for example, traffic growth.

One very important thing that the above barriers have in common is that they are not 100% effective (individually or collectively) even when operating fully to specification. The degree and extent to which the barriers are able to reduce risk (by removing conflicts) depends, in the first place, on the functionality and performance of the various elements of the ATM system that underlie each barrier. Thus the ATM system needs to be designed such that, when all the barriers operate as intended, the aggregate reduction in risk is more than enough to achieve a tolerable level of safety[13] - “more than enough”, of course, in order to provide a risk margin to account for occasional failure of the barriers.

It is from this perspective, and from experience on recent EUROCONTROL ATM programmes, that the new framework for safety assessments, outlined in the next section, was developed.

A New Framework

The recommended starting point for any

[9] The idea of barriers comes from Professor James Reason’s well-known “Swiss Cheese” model of safety management.

[10] Under some circumstances, the operation of the barriers can overlap in time.

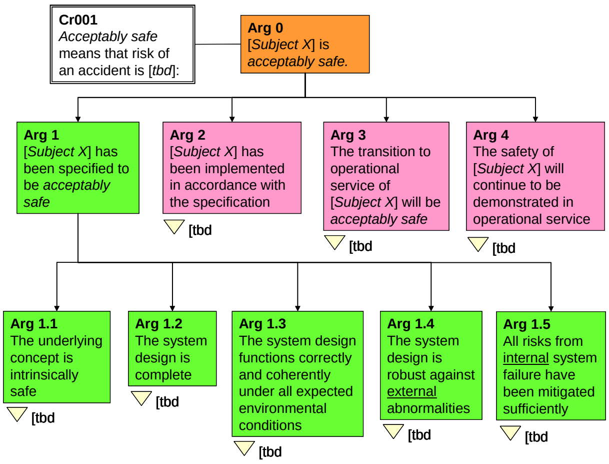
[11] The term Conflict is used herein according to the definition in [7] – ie “any situation … in which the applicable separation minima may be compromised [infringed]”.

[12] Depending on the particular implementation of STCA

[13] Note that in ECAC airspace, achievement of the tolerable level of safety (specified by ESARR 4) has to be demonstrated without taking account of the benefits of (Collision Avoidance) safety nets – see [8].

safety assessment (see [9]), is the production of a high-level, structured Safety Argument.

The new framework for ATM safety assessments follows that recommendation, as illustrated in Figure 4, using a simplified form of Goal-structuring Notation (GSN). The proposed Argument is intended to be generic and capable of adaptation to a wide range of applications, covering both the safety assessment of an on-going ATM service and a proposed change to an ATM service / system[14].



Cr001
Acceptably safe means that risk of an accident is [tbd]:
Arg 0
[Subject X] is acceptably safe.
Arg 1
[Subject X] has been specified to be acceptably safe
Arg 2
[Subject X] has been implemented in accordance with the specification
[tbd
Arg 3
The transition to operational service of [Subject X] will be acceptably safe
[tbd
Arg 4
The safety of [Subject X] will continue to be demonstrated in operational service
[tbd
Arg 1.1
The underlying concept is intrinsically safe
[tbd
Arg 1.2
The system design is complete
[tbd
Arg 1.3
The system design functions correctly and coherently under all expected environmental conditions
[tbd
Arg 1.4
The system design is robust against external abnormalities
[tbd
Arg 1.5
All risks from internal system failure have been mitigated sufficiently
[tbd

Figure 4 – Safety Argument

The top-level Claim (Arg0) states that the subject (ie on-going service or change) is *acceptably safe.* Strictly speaking, in the case of a change, this would be an abbreviated way of saying that the ATM service, following the introduction of the change, is acceptably safe.

What is meant by acceptably safe in Argument 0 is defined by the safety criteria in Cr001 – these may be defined:

- absolutely - eg compliance with a Target Level of Safety (TLS), and/or
- relatively - eg risk to be no higher than, or (where a safety improvement is required) substantially lower than, the pre-change situation, and/or
- minimally – eg risk to be reduced as far as reasonably practicable (AFARP).

The Claim is then decomposed into four principal Safety Arguments, using the GSN convention that an Argument can be considered to be true, if (and only if) each of its immediate 'offspring' can be shown to be true[15].

Arguments 2 to 4 reflect normal ATM safety practice and are not expanded herein – for further guidance see [9]. However, it is important to note that Argument 1 applies to the Concept as a whole; therefore, where such concepts are implemented in stages, the term "transition" in Argument 3 should be interpreted as including the safety of each stage of a phased deployment of the end system.

It is the decomposition of Argument 1 which reflects the Success Approach (Arg1.1 to Arg1.3) and Failure Approach (Arg1.4 and Arg1.5)[16]. Typical issues to be addressed under each Argument are discussed in the rest of this section – in each case, the human, procedure, equipment and airspace elements of the system must be considered.

Intrinsic Safety of the Concept (Arg1.1)

Need to show, inter alia, that:

- the Concept of Operations fully describes how the system is intended to operate
- differences from existing operations have been described, understood and reconciled
- the impact of the concept on the operational environment has been assessed and shown to be consistent with the safety criteria
- the key functionality and performance parameters have been defined and shown to be consistent with the safety criteria.

The issues here are whether the basic idea is intrinsically safe – ie whether the Concept is capable of satisfying the safety criteria, assuming that a suitable system design could be produced – and what the key parameters are that make it so.

Design Completeness (Arg1.2)

Need to show that:

- the boundaries of the system are clearly defined
- everything necessary to achieve a safe implementation of the Concept – related to equipment, people, procedures and airspace design - has been specified (as safety requirements)
- all safety requirements on, and assumptions

[14] System includes airspace design, equipment, people and procedures

[15] At the lowest eventual level of decomposition, of course, an Argument can be considered to be true if there is adequate Evidence to show that it is.

[16] It is a moot point whether Arg1.4 should be in the success approach or the failure approach. In practice, the distinction between the *success* and *failure* approaches is unimportant compared with ensuring overall that everything required by Arg 1.1 to 1.5 is covered

about, external[17] elements of the end-to-end system have been captured

- safety requirements are realistic – ie are capable of being satisfied in a typical implementation in hardware, software, people and procedures.

The main question here is whether we have thought of everything, in terms of the design, that is necessary to fully implement the Concept.

Design Correctness (Arg1.3)

Need to show that:

- the design is internally coherent – eg consistent functionality and use of data throughout the system
- all reasonably foreseeable normal operational conditions / range of inputs from adjacent systems have been identified
- the design is capable of delivering the required risk reduction under all reasonably foreseeable normal operational conditions / range of inputs
- the design functions correctly in a dynamic sense, under all reasonably foreseeable normal operational conditions / range of inputs

The main question here is whether we have maximised the opportunity to reduce risk, over the full range of conditions that the system is likely to be subjected to in its operational environment.

Design Robustness (Arg1.4)

Need to show that:

- the system can react safely to all reasonably foreseeable external failures – ie any failures in its environment / adjacent systems that are not covered under Arg1.3
- the system can react safely to all other reasonably foreseeable abnormal conditions in its environment / adjacent systems.

Here we are concerned with abnormal conditions in the operational environment, from two perspectives: can the system continue to operate effectively – ie reduce risk; and could such conditions cause the system to behave in a way that could actually induce a risk that would otherwise not have arisen?

Mitigation of Internal Failures (Arg1.5)

This relates to the more 'traditional' failure-based approach to ATM safety assessment. Unlike Arguments 1.1 to 1.4, which lead to a specification of the risk-reducing properties of the system (ie safety requirements for the functionality and performance of the system), Argument 1.5 leads mainly to a specification of Safety Objectives[18] and Safety Requirements for the integrity of the system.

Typically, need to show that:

- All reasonably foreseeable hazards, at the boundary of the system, have been identified
- The severity of the effects from each hazard has been correctly assessed, taking account of any mitigations that may be available / could be provided external to the system
- Safety Objectives have been set for each hazard such that the corresponding aggregate risk is within the specified safety criteria
- All reasonably foreseeable causes of each hazard have been identified
- Safety Requirements have been specified (or Assumptions stated) for the causes of each hazard, taking account of any mitigations that are / could be available internal to the system, such that the Safety Objectives are satisfied
- those Safety Requirements are capable of being satisfied in a typical implementation in hardware, software, people and procedures.

Here we are concerned with the internal behaviour of the system, from two perspectives: how loss of functionality could reduce the effectiveness of the system in reducing risk; and how anomalous behaviour of the system could induce a risk that would otherwise not have arisen.

Further guidance on the application of the above framework is given in the next section of the paper, in the form of examples drawn from some recent applications of it in EUROCONTROL. More detailed guidance on the *failure* approach in particular may be obtained from [5] and [9].

Example Applications

This section gives a brief outline of how the above safety assessment framework has been (or could be) applied to three current European safety assessments.

In each case, a short introduction is followed by the specified safety criteria and a description of

[17] The term 'external' here usually refers to those elements that lie outside the managerial control of the organisation accountable for the safety assessment

[18] Safety Objectives is a term used in ESARR 4 [3] and the EUROCONTROL Safety Assessment Methodology [6] to describe the maximum tolerable occurrence rate of hazards.

what work is involved in addressing each of the five main branches of the Safety Argument[19].

EUR RVSM

The introduction of RVSM, between FL290 and FL410, in January 2002 was heralded as the biggest change in EUR airspace for more than 50 years. It required all 41 States involved to implement the change at precisely the same time, having first obtained the approval of their respective safety regulatory authorities.

RVSM Safety Criteria

Overall, EUR RVSM is required to satisfy three safety criteria:

- the ICAO TLS of $\leq 5\text{x}10^{-9}$ accidents per flight hour (pfh), including a failure-free component, due to aircraft technical height keeping error, of $\leq 2.5\text{x}10^{-9}$ accidents pfh
- the post-RVSM accident rate to be no greater than the pre-RVSM rate
- the risks associated with RVSM to be reduced as far as reasonably practicable (AFARP).

Intrinsic Safety of RVSM Concept

The decision, in the 1960s, to set vertical separation above FL290 at 2000ft was based on the concerns about baro-altimetry accuracy at these higher altitudes. Clearly, the fundamental (intrinsic) safety of RVSM necessarily depends on modern altimetry and autopilot systems being able to maintain aircraft at their assigned altitude to an accuracy commensurate with 1000ft vertical separation.

Key Functional Safety Requirements for the aircraft equipment are specified in RVSM Minimum Aircraft System Performance Specification (MASPS). On-going proof of compliance with these requirements in the EUR Region is the subject of a major height-monitoring programme (and associated collision-risk modelling exercise) involving five Height Monitoring Units positioned at key points around Europe.

In terms of the possible effects of RVSM on the safety of the operational environment, a number of issues were considered, including the effect on:

- the pre-existing risk associated with “level busts”
- RVSM-incompatible versions of TCAS (V6.04a)
- RVSM-compatible versions of TCAS (V7.0) in terms of rate of Nuisance alerts
- the severity of Wake Vortex and Mountain Wave encounters.

RVSM Design Completeness

The design of the system to support RVSM covered the following main areas, for which Functional Safety Requirements were specified:

- Airspace Design – eg FL orientation, RVSM / CVSM transition areas, and resectorisation
- Flight Crew Procedures and Training – eg aircraft operational procedures, RT phraseology
- Aircraft Equipment – see above
- ATC Procedures and Training - eg ATC operational procedures, RT phraseology
- ATC Equipment – eg display of RVSM status, modification of STCA parameters
- Flight Planning – including AOs and IFPS
- System Monitoring – eg MASPS compliance, operational errors, collision-risk assessment

RVSM Design Correctness

Proving the correctness and coherency of the design of the EUR RVSM system was based on:

- about four years prior operational experience of RVSM in the NAT Region
- a five-year programme of fast- and real-time simulations, in 11 key areas of EUR airspace

RVSM Design Robustness

Assessment of the robustness of the design of the EUR RVSM led to the development of additional Flight Crew and ATC procedures (and associated training) for, inter alia: reporting and handling of aircraft emergencies, lost comms and loss of RVSM capability.

Mitigation of RVSM Internal Failures

This followed a ‘conventional’ safety assessment approach, and included analysis of, inter alia: initial flight-planning errors, Flight Crew operational errors, ATC operational errors, aircraft equipment failures, ATC equipment failures.

Time-Based Separation

Time-Based Separation (TBS) is a new concept of using aircraft separation based on time intervals for landing in strong headwind conditions. The problem with the currently used Distance-Base Separation (DBS) is that the time taken to cover the distance intervals between aircraft increases as

[19] Strictly speaking, the framework described herein was developed as a result of the work done on the examples provided. Therefore, although most of the examples complied with the principles of the success and failure approach, they did not necessarily follow the precise form of Safety Argument proposed in this paper.

aircraft ground speed decreases, resulting in decreased runway capacity during periods of strong headwinds. The objective of the on-going EUROCONTROL TBS project is to investigate the possibility of recovering any such loss of runway arrival capacity, at busy airports, while maintaining required levels of safety.

TBS Safety Criterion

A relative approach is being taken in the safety assessment of the TBS concept. TBS will be considered to be acceptably safe if it can be shown that the risk associated with TBS scenarios is no higher (and preferably lower) than for the equivalent DBS scenarios.

Intrinsic Safety of TBS Concept

In order to avoid loss of runway arrival capacity, TBS (time) minima have to be no greater than the time interval that would exist if DBS minima were applied in zero-wind conditions – ie the minimum distance between aircraft under TBS are reduced, in comparison to DBS distances, in proportion to the strength of the headwind.

However, the DBS (distance) minima themselves have to take account of two key safety considerations:

- the Wake Vortex Encounter (WVE) risks during normal operations - ie separations and operating conditions are as designed, and there have been no system failures
- the mid-air collision (MAC) risks due to limitations in surveillance radar performance – particularly accuracy and resolution.

Therefore, we need to be sure that the reduction in separation distances between aircraft, which results from TBS, does not increase either of the above risks.

The WVE issue is complex since wake vortex effects generally decay with time, reduce with distance from the generator aircraft, and dissipate more quickly in rougher air conditions. It will be necessary, therefore, to carry out Wake Vortex Encounter modelling in order to assess the relative risks (TBS cf DBS) and set the TBS separation minima to meet the safety criterion – ie such that TBS risks of encountering a vortex of a certain circulation speed are no higher than for DBS.

If TBS leads to separation minima below the current radar-defined minima, new (safety) requirements for radar surveillance will be specified such that the current risk of MAC is not exceeded.

The effect of TBS on the operation of safety nets – particularly STCA - will also need to be considered. The smaller average aircraft distance separations under TBS may limit the effectiveness of STCA, unless STCA is suitably modified.

TBS Design Completeness

Issues to be considered in this context include:

- Procedures for determining when and how to apply TBS rather than DBS[20],
- Procedures for applying TBS for specific WVE cases – eg light aircraft following heavy aircraft
- requirements for ATC Support Tools to calculate the aircraft spacing necessary to achieve the minimum time separations accurately
- ATC Display requirements
- ATCO Training in TBS procedures.

TBS Design Correctness

Issues that will need to be considered include:

- effect on ATCO workload and effectiveness
- the effects of changeover from DBS to TBS and vice versa
- interfacing / coordination between TBS and DBS airspace - conceptually, DBS may continue to be applied in a TBS environment for all flight phases prior to final approach intercept (or to a limited area in the vicinity of the intercept)
- interactions between TBS and other APP / TWR procedures, including the need to protect the OFZ and (where applicable) ILS Localiser Sensitive Area.

Real-time simulations will be an important part of assessing the dynamic behaviour of TBS.

TBS Design Robustness

There are least three key issues here:

- the possibility of sudden changes in wind conditions. In the required WVE risk modelling, the resulting WVE probability/ severity curves will be factored by the estimated occurrence frequency of this event
- the increased interdependencies between the arrival manager (AMAN) and TBS - the former is the source of the information for Trailing Target Positions (TTP) - a support tool which displays the minimum time separation between aircraft
- The effects of variation in actual aircraft groundspeed, due to TBS being based upon nominal groundspeed values.

[20] In principle, application of minimum time-based separation intervals on final approach , as opposed to minimum distance-based radar separation minima, could also be applied consistently across all wind conditions.

Mitigation of TBS Internal Failures

The failure-risk assessment has not yet been completed. Potential failures that need to be considered, and addressed in the design of the Controller Support Tool(s), include:

- Incorrect calculation of TBS minima, by ATC Support Tools.
- Incorrect application of TBS minima, by ATCO
- Pilot non-compliance with ATC instructions. This refers to the effect of TBS on pilot situational awareness, as a result of which pilots may disbelieve ATC due to the resulting unfamiliarly close separation.

ACAS II

The Airborne Collision Avoidance System (ACAS) II is an airborne safety net designed to provide a reduction in the risk of mid-air collision. In 1995, a common ACAS policy and implementation schedule for the mandatory carriage of ACAS II in EUR airspace was agreed. The final date for compliance with the Phase 2 requirements of the Mandate was 31 March 2006.

In addition to extensive ACAS safety assessments which have been already undertaken, EUROCONTROL is conducting a post-operational safety assessment of ACAS II in EUR airspace.

ACAS II Safety Criterion

ACAS II supports the Collision Avoidance barrier (see Figure 3 above). As a safety net, its functionality and performance is designed to provide a significant reduction in the residual risk of collision, resulting from the inherent limitations or failure of the preceding barriers. However, it also carries with it the possibility of system failure and adverse side effects arising from its operation, both of which have the potential to erode its safety benefit.

In accordance with EUROCONTROL policy [8], the safety benefit of ACAS II may not be counted in demonstrating compliance with the ESARR4 safety target [3]. Therefore, the ACAS II safety assessment uses a relative safety criterion - that the risk of an accident (from the aggregate effects of risk reduction by, and risk increase due to adverse side effects and failure of, ACAS II) shall be substantially lower than without ACAS II. In effect, ACAS exhibits the general characteristics of a *secondary* SRS illustrated in Figure 2 above.

Intrinsic Safety of ACAS II Concept

The intrinsic risk-reducing capabilities of ACAS II have been established by means of so-called Risk Ratio modelling which computes the effectiveness of ACAS II in resolving mid-air collisions based upon an encounter model, the ACAS II algorithms, and other factors. The results of this modelling predict that, even taking into account those known factors which reduce ACAS II effectiveness in normal operation, there is a significant reduction in the risk of collision.

ACAS II Design Completeness

The ACAS II system is specified via operational and equipment requirements and standards[21] which need to be complied with by ANSPs, airspace users, equipment manufacturers, and certification authorities. In turn, these requirements have been subject to local transposition and implementation by these parties in order to allow ACAS II to be deployed consistently throughout EUR airspace. Due to the wide scope of ACAS II, this raises issues which are being addressed within the safety assessment, including:

- demonstrating that the scope of the requirements and standards are consistent with the ACAS II Risk Ratio models
- determining whether the various ACAS II requirements and standards are sufficient to describe ACAS II for the purposes of its stakeholders
- determining the degree to which transposition and implementation of the requirements and standards by stakeholders should be directly demonstrated by evidence, rather then being inferred
- demonstrating that available stakeholders' evidence is consistent with the scope of the corresponding ACAS II requirements and standards
- demonstrating that the interactions with, and assumptions about, systems which interface with ACAS II have been captured at all levels of the design.

ACAS II Design Correctness

Assurance of the correctness of the ACAS II specifications is being addressed primarily by:

- establishing that functional consistency exists between the Risk Ratio modelling, ACAS II requirements and standards, and equipment 'design' documentation
- confirming the adequacy of ACAS II procedures, training and equipment by operational monitoring and effecting corrective actions as required.

[21] ICAO, RTCA, FAA, *et al*

ACAS II Design Robustness

Design robustness is addressed in part by an event tree in the Risk Ratio model. This allows the influence of various 'real-world' factors on the theoretical effectiveness of ACAS II logic to be established quantitatively. These factors include;

- collision geometry and visual acquisition
- transponder/ACAS equipage and its abnormal behaviour
- presence of an ATC service
- controller and pilot possible behaviours surrounding the RA event.

These factors equate to abnormal situations which may occur within different parts of the overall ATM system. The ACAS II safety assessment will determine whether this is a sufficient assessment of robustness to cover all levels of ACAS II design.

Mitigation of ACAS II Internal Failures

Since the development of ACAS II systems pre-dated ESARR 4 [3], it has not so far benefited from a systematic *a priori* identification and assessment of hazards in accordance with the SAM methodology [5]. Such hazards could conceivably arise from undesirable side effects of ACAS II in normal operation such as, for example, a preponderance of nuisance alerts. Moreover, there will be hazards arising from system failure; false alerts for example. Any increase in risk from such hazards needs to be argued as being negligibly small. Both these types of hazards would be considered as belonging to the failure approach because they are risk-increasing, even though the former type arises during normal operation.

The systematic identification, assessment and mitigation of ACAS II hazards (beyond those already detected during ACAS II operational monitoring) will form the final stage of the ACAS II safety assessment.

Conclusions

The work that is the subject of this paper was initiated by concerns that, in the past, safety in ATM has been considered to be synonymous with reliability / integrity and that, consequently, safety assessments have often been far too narrow in their scope.

It is argued that ATM is a very effective way of reducing risk inherent in aircraft operations and that its effectiveness greatly outweighs any new risks that failures of the ATM system might induce. Therefore, in assessing the safety of new concepts in ATM we must be at last concerned with assuring the continuing effectiveness of ATM (the *success* approach) as we are with potential failures in the implemented ATM system (the *failure* approach)

The idea behind this broader approach to safety is captured in a high-level, generic safety argument framework that has been developed from experience on a number of recent ATM development programmes.

It has proved difficult, in the context of this paper, to develop the idea further into a "one fits all" detailed framework, since the greater detail can rapidly become concept specific. Instead, the paper has used simplified applications of the framework to three recent / current ATM programmes to illustrate the importance of the success approach and to give clues as to how the framework could be developed for other applications.

Having demonstrated the effectiveness of the approach on relatively simple applications, the next big challenge will be to develop the ideas in a way that will be appropriate to, and adequate to support, much bigger, future ATM concept developments, including those associated with SESAR.

References

[1] Leveson N G, 2001, *The Role of Software in Recent Aerospace Accidents*, 19th International System Safety Conference, Huntsville AL, USA,

[2] Krastev A, Smith B, Masson M and den Hertog R, 2006, *Preliminary Hazard Analysis of EUROCONTROL Concept of Operations 2011 using the FAST Method*, EUROCONTROL Safety R&D Seminar, Barcelona, Spain

[3] EUROCONTROL, 2001, *ESARR 4 - Risk Assessment and Mitigation in ATM*, Ed 1.0

[4] EUROCONTROL, 2004, *EUR RVSM Post-Implementation Safety Case*, Ed 2.0

[5] EUROCONTROL, 2006, *Air Navigation System Safety Assessment Methodology*, Ed 2.1

[6] IEC, 2000, *IEC 61508 - Functional Safety of Electrical/Electronic[etc] Safety Related Systems*,

[7] ICAO, 2005, *Document 9854 - Global Air Traffic Management Operational Concept*, Ed 1

[8] EUROCONTROL, 2003, *SRC Policy Document 2, Use of Safety Nets in Risk Assessment & Mitigation in ATM*, Ed 1.0

[9] EUROCONTROL, 2006, *Safety Case Development Manual*, V 2.2

[10] Hammer J, Caligaris G, Llobet M, 2007, *Safety Analysis Methodologies for ADS-B-based*

Surveillance Applications, 7th USA/Europe ATM R&D Seminar, Barcelona, Spain

Biographies

Derek Fowler has over 30 years aerospace engineering experience. Since 1990 he has specialised mainly in ATM systems, for the last 9 years acting as a safety consultant. He is currently working under contract for EUROCONTROL, providing safety expertise to ATM programmes.

Gilles Le Galo has worked as an ATCO and OJTI for 20 years. Since 1990 he has specialised in safety management and currently works at EUROCONTROL HQ as SMS Implementation Coordinator for ECAC ANSPs. He is the project manager for the 'Support to ANSPs for SMS Implementation' (SASI) project.

Eric Perrin has twelve years' experience in aviation, six of which have been spent on safety assessment and management. Currently he is Deputy Safety Research Team Manager and Safety Management System (SMS) Manager at the EUROCONTROL Experimental Centre.

Stephen Thomas has worked in aviation safety for 30 years, the last nine specializing in ATM. He is an independent safety consultant and has worked with EUROCONTROL since 2001 on Safety Management Systems and safety assessments.

Acknowledgements

The authors would like to acknowledge the invaluable contribution made by Patrick Mana, of EUROCONTROL, in the development and review of this paper.

The views expressed herein are authors' own and do not necessarily reflect EUROCONTROL policy.