Safety Science, Volume 42, Issue 8, October 2004, Pages 723-754.

Cranfield University
Building 83
Cranfield
Bedfordshire MK43 0AL
England
Tel +44 (0) 1234 750111 Extn.5086
Fax +44 (0) 1234 750192
e-mail: p.brooker@cranfield.ac.uk

AIRBORNE SEPARATION ASSURANCE SYSTEMS: TOWARDS A WORK PROGRAMME TO PROVE SAFETY

Peter Brooker

[Cranfield Research Report PB/12/1/02]

[ISBN 1 861940 91 2]

"Content is a glimpse of something, an encounter like a flash. It is very tiny – very tiny, content." Attributed to Willem de Kooning.

ABSTRACT

With Full Delegation Airborne Separation Assurance System (ASAS), separation control would be delegated to the (properly equipped) aircraft, i.e. aircraft pilots are responsible for aircraft separation. The aim is to try to identify a tangible work programme – rational and evidence based, and within the compass of known techniques – that would prove safety. The task here is to create a framework in which to integrate these existing building blocks with results from additional work developed from well-specified experiments.

Reasons for retaining the existing separation minima in an ASAS system are put forward. For the current system, comparatively large proportions of the Air Traffic Services risk budget should be allocated to 'Reasonable Intent' risk (effectively 'right place on wrong flight path'). The key argument here is that mid-air collision in an ASAS environment will predominantly arise from this type of risk. The use of Probabilistic Risk Assessment, which requires the probabilities of safety-critical events to be estimated for 'human components' (Human Reliability Analysis), is reviewed. The danger is the creation of 'over-elaborate' models – ones whose parameters cannot be reliably estimated from the data

likely to be obtainable. A simple model that can be soundly based on available data is proposed.

1. INTRODUCTION

The present air traffic control (ATC) system has several distinct components. Air traffic controllers communicate through radiotelephony; they use flight plans agreed with pilots; they monitor highly processed radar data; and, in developed States, they have short-term conflict alert (STCA) systems available to warn of aircraft coming into close proximity. These data flows are embedded in 'safety structures', e.g. with well-defined controlled airspace and formal rules for control such as the minimum separation permitted between aircraft. System procedures were originally designed to be 'tolerant' of equipment failure, as decades ago the equipment was much less accurate or reliable. The system safety structures have also been designed to be relatively easy for human operators to comprehend and use.

The present concept has evolved: it is 'overlaid', in that new technology has largely been added on to the previous concept. Any new functionality generally has to able to carry out both the tasks of the previous generation plus some new ones. Can the current system continue 'evolving' for many more years? To quote Amalberti (2001):

> "Most of today's man-machine systems were designed in the 1960s…No system will last forever, and we are probably dealing today with ageing logic which will someday be replaced by a different logic once the technology is finally mature…"

[The phrase 'once the technology is finally mature' is a key one, and is the subject of much of the following.] There are already major problems with system capacity, leading to significant flight delays, mainly because controller workload is nearing acceptable limits in busy airspace sectors. The phrase 'Free Flight' has resonated in the ATC community for much of the last decade, suggesting that flights should use preferred – more cost effective – routeings and profiles, rather than follow fixed routeings.

Airborne Separation Assurance Systems – ASAS – are a possible way forward for air traffic management. Studies are in progress under the auspices of the FAA, Eurocontrol (FAA/Eurocontrol, 2001) and the European Commission. Some technical details of potential technologies to provide ASAS, and the wider system context, are given in EC (1998) – ASAS would generally be provided through some kind of ADS-B ('Automatic Dependent Surveillance – Broadcast') equipment. The main feature of ASAS is that separation control is delegated to the (properly equipped) aircraft, i.e. aircraft pilots are responsible for their aircraft's separation from other flights. This is termed 'full delegation', defined as:

> **Full delegation:** Pilots are responsible for all the tasks related to separation assurance: identification of problems and solutions, implementation and monitoring.

ASAS as used here always refers to Full Delegation. There are other ways in which airborne equipment providing information on nearby aircraft can be integrated into the system, e.g. see Brooker (2003). There might be critical – and potentially very difficult – transition processes between the present system and ASAS. Such transitions might raise new problems both over time – e.g. mixed equipage and partial delegation of ATC tasks – and in space – e.g. interfaces between ASAS and conventional ATC regions. However, intermediate operational steps to the 'full ASAS' investigated here could provide useful evidence on ASAS safety performance. No assumption about the availability of such information is made here, i.e. it would be a 'bonus'. Some of the ideas and techniques explored here could also have wider applicability; for example, the development of automated aids for ground control would produce similar issues to those raised by ASAS.

ASAS would be considerably different from the present Air Traffic Management (ATM) system (taken here to include ATC and all other functions that deliver air traffic's safety, capacity, cost-effectiveness, etc). This paper addresses the question: "How would one prove that an ASAS system is safe in practice?" The biggest problem with safety arguments is the need to understand all the possible ways that a failure could lead to an accident. Demonstrating completeness is extremely difficult with complex systems such as ASAS in busy airspace – or anything else that relies heavily on software, displays and people. There is inevitably some resort to probabilistic arguments and statistical evidence. So, the question addressed here might better be put in the form: "What would be the least complex and most robust calculations required to do the job?" It needs to be stressed that broad-based real-time simulations are inadequate to demonstrate safety of complex systems, but, as will be indicated in later sections, well-focused simulations are indispensable components of effective risk analyses.

ASAS has, of course, to be seen in the larger context of possible future ATM systems – as is evident from the linkages from the European Commission's THEATRE work. [THEATRE is the 'Thematic Network on Air Transport for ATM Validation Activities': its objective is to achieve transparency and effectiveness between validation projects in the 5th Framework Programme of the European Commission; its Validation and Safety Working Group is particularly relevant here.]

The validation problem can be stated in even more demanding terms. To quote the recent FAA/Eurocontrol document on ASAS (2001):

> "ASAS applications involving major reliance on aircraft systems and changes to present responsibilities and procedures to ensure aircraft separation will require rigorous safety analysis and validation before

> implementation. This analysis will need to demonstrate conclusively that the ASAS application meets or exceeds the required Target Level of Safety, including consideration of equipment failure and human error. Methodologies and guidelines for these analyses will need to be agreed at the international level."

Words such as 'rigorous' and 'conclusively' set a very uncompromising tone. The same kinds of phrases appear in many other ASAS-related documents – 'detailed and rigorous safety assessment' is a common variant. But the validation processes then described often appear to be very abstract and unspecific: indeed, some appear to value metaphysics and elegance as having more merit than practical applicability. So, what would actually be practical and theoretically sound methods for achieving the quantitative goal? Precisely what calculations have to be carried out? What evidence and information from the operation of present systems needs to be used? What new data has to be gathered – and does this require specific types of simulations? How, in Amalberti's phrase, is the technology – or rather the technology & human system – demonstrated to be 'fully mature'?

This paper attempts to start to answer such questions. The aim is to try to identify a tangible work programme – rational and evidence based, and within the compass of known techniques – that would prove safety in the terms set out in the above quotation. In fact, many of the building blocks already exist: the task here is to create a framework in which to integrate these components with additional work that can be developed from well-specified experiments.

The sections are as follows:

2. The Concepts of Validation and Proof
3. Safety Targets and Probability
4. Separation Minima
5. Probabilistic Risk Assessment
6. 'Reasonable Intent' Risks
7. ASAS Safety Issues
8. A 'Minimal Framework' Collision Risk Model
9. Conclusions

2. THE CONCEPTS OF VALIDATION AND PROOF

It is necessary to try to understand the logical underpinnings of 'validation'. ATM has a definition (Eurocontrol, 1998):

> **Validation –** 'The process through which a desired level of confidence in the ability of a deliverable to operate in a real-life environment may be

> demonstrated against a pre-defined level of functionality, operability and performance.'

This is a complex definition. It has moved on from the sort of definition that is given in dictionaries (Chambers, 1998) eg:

> **validate** -…to check items to ensure that they conform to input rules and e.g. fall within an acceptable range

Validation in ATM appears to be being used to mean 'proof'; again from Chambers:

> **proof** -…demonstration; evidence that convinces the mind and goes toward determining the decision…;

The key phrase above that supports this view is 'desired level of confidence', which equates to 'evidence that convinces the mind'. One reason that the word proof is not used may be that the same or similar word is found in several European languages, where it tends to have a meaning of 'test'. To understand how proof should be manifested in ATM, it is worth examining the concept in logic and legal contexts, as these influence thinking about the concept.

Proof in logic tends to have a very restricted meaning. It is a demonstration of the validity of a proposition based on specific premises – a deductive argument. Thus, for any real number p, $1 + p^2 \geq 2p$ can be shown to be correct given the rules of algebra and the fact that the square of a real number is positive. Engineering design problems such as ASAS do not fall into this class, because they rely on an 'inductive' chain of reasoning, in which the truth of the premises makes it probable – rather than necessary – that the conclusion is true. Hence, 'proof' of ASAS safety cannot be logically guaranteed – an obvious point, but one at odds with calls for 'absolute safety'. In other words, it is impossible to assure any sort of formal 'completeness' of any risk analysis – there can never be a proof that all types of errors or risk modes have been identified. However, in traditional 'hard' engineering disciplines, with physical constructions, extrapolations based on measured data combined with physical laws can enable the performance of a system to be proved satisfactory to a high degree of confidence.

When human beings enter into the analysis, either as part of the operational system or as analysts of its performance, the situation is much more complex. Issues raised by the former are dealt with in latter sections. An example of the latter is expert witnesses giving scientific – generally medical – evidence in a legal case. They have two tasks: to provide basic scientific or technical data; and to present conclusions or inferences from the facts, given that the judge and/or jury, not having specialised knowledge, could not themselves draw. In principle, such medical opinions can be empirically supported, although the required range of observations and experiments may not be available.

In an adversarial legal system, such as the UK and USA, evidence from expert witnesses can be challenged, and indeed both sides in a legal case can present expert testimony. The sort of question with which the expert witness has to assist is “Has the prosecution proved beyond reasonable doubt that the accused killed the victim?” These are obviously similar to the phrases ‘desired level of confidence’, and ‘evidence that convinces the mind’. Decision-makers cannot be expected to accept the expert’s judgements and opinions in an uncritical fashion. There is an onus on him or her to be able to demonstrate to these intelligent non-experts the reasoning and processes by which the conclusions have been reached. This demonstration should be ‘robust’, i.e. with the inferences not easily being demolished if particular assumptions do not fully hold or relationships between variables are not exact. The decision-makers – States, ATC organisations, and individuals – are personified here as the ‘Rational Evidence Scrutiniser’ (RES). The RES represents the ‘directing minds’ responsible for aviation safety.

Given the definition of validation set out earlier, the task of proof would need to be a convincing description explaining how safety is assured through protective barriers in the system and the nature of resilience against ‘unsafe incidents’. This narrative would have to make clear the nature and rationality of the quantifications involved. Confidence is built by displaying with clarity the rational nature of the calculations and demonstrating that all the quantitative aspects have some reasonable basis in observations and relevant measurements. This ‘hard’ viewpoint stretches back to Descartes and Hume. What else could be adopted in critical matters of aviation safety, because such processes must reflect the highest standards of public decision-making?

The RES therefore wants to be convinced through a rigorous and explicit approach that risk estimates are well founded. It is the actual ‘process of explication’ that should convince the RES that there are firm foundations to risk estimates. One would expect the RES to be sceptical if the core arguments were to be attempted through scientific complexity, with ‘magic black box’ mathematical and statistical calculations involving many acronyms and a host of Greek symbols. ‘Confidence’ could surely only come about through the creation of a compelling ‘narrative’ explanation, soundly based in theoretical understanding and empirical evidence, and open to challenge, checking and verification at every stage. In particular, this implies that commercial modelling products, i.e. which are not open to the possibility of scrutiny or peer review, would not be adequate to convince the RES.

It needs to be noted that validation is not the only critical element in ATM safety processes. Certification and Verification are key elements:

> **Certification –** ‘The process aiming at the satisfaction of an authority that a deliverable complies with a set of regulations, in order to ensure its proper operation.’

> **Verification –** `The process of evaluating the products of a given system development activity to determine correctness and consistency with respect to the products and standards provided as input to that activity.'

These will not be discussed directly in the following, although many similar issues are involved – in particular, safety certification increasingly relies on the outputs from validation and safety management.

More generally, safety criticality of ASAS applications is being explored by a number of researchers, mainly with aim of guiding developers and designers to ensure robustness where it is most needed. The papers by Zeitlin (2001) and Zeitlin and Bonnemaison (2000), and their references to RTCA work, are particularly relevant to some of the issues discussed here.

3. SAFETY TARGETS AND PROBABILITY

The key quantitative safety concept in ATC is that of a Target Level of Safety – TLS. This is actually a design hurdle. It is a quantified risk level (measured as an accident rate) that a system should – i.e. be designed to – deliver. TLSs cover all aviation-related causes, but do not usually attempt to cover the consequences of terrorism or criminal behaviour (although the literature has not always been clear on this). It is usually expressed as a proportion of fatal accidents per so many flying hours (or airport movements when that is more appropriate). As will be examined later, most of the practical problems are not actually with the TLS but with the proper estimation of the safety level that is – or would be – achieved. There is an Actual Level of Safety – an 'ALS' - being achieved in the system under examination: how is this to be calculated with sufficient accuracy for the RES to be confident that the ALS < TLS?

A TLS can be derived in several ways. Brooker (2002) and the Safety Regulation Commission of Eurocontrol (SRCb, 2000) sketch the kinds of calculations involved. TLSs appropriate for accidents arising from mid-air collisions have been developed since the 1970s. They are usually derived by taking historical accident rates, which show a progressive reduction over time, and extrapolating forward. Thus, the TLS value gets tighter and tighter over time. The original focus was on commercial passenger jet flights in North Atlantic airspace, but the TLS has been used for en route controlled airspace generally. In particular, it is used in the calculation of the separation minima required between aircraft in oceanic and domestic airspace.

The TLS is measured in fatal aircraft accidents, i.e. accidents in which at least one person in the aircraft was killed, per so many aircraft flying hours. The current ICAO (RGCSP, 1995) figure of 1.5×10^{-8} fatal aircraft accidents per flying hour is the total rate corresponding to mid-air collisions – for any reason and in any spatial dimension – in en route flight in controlled airspace. Brooker and Ingham (1977), and Davies and Sharpe (1993) show how the TLS is derived. It is important to stress that the TLS includes the consequences of 'blunder' type errors, such as errors in coordination between the aircraft crew and ATC (leading to an aircraft occupying a flight level or routeing other than that intended by ATC), or errors in ATC instructions leading to a similar consequence. For example, Harrison and Moek (1992) explain how the vertical domain TLS, taken as third of the total TLS, is partitioned into 'loss of planned separation' errors and 'other' errors. Thus, the TLS is not simply driven by technical operation – e.g. altimetry in the vertical case – but by total system performance. This is a crucial point, which underpins the discussion in Section 6, on 'Reasonable Intent' risks.

To put the ICAO TLS in context, for UK ATC, assume 1 million (10^6) en route flight hours a year indefinitely into the future. If the TLS represents the actual risk rate, this would correspond to one mid-air collision – two fatal aircraft accidents – per 134 years. Given current average passengers per aircraft, there would be about 200 fatalities in such a collision. On past decision-making trends, such a TLS for 20 and 30 years ahead with an ASAS-based ATC will be even tougher. Such a target will requires the acceptance of the aviation community, given that ASAS as described here would in many ways be an 'end state' concept for ATC.

But what does such a design target mean in practice? To explore this, it is necessary to examine the statistics of rare events.

Poisson Distributions and Confidence Intervals

The Poisson distribution – found in any standard textbook on probability – is a good statistical model for discrete and rare events, particularly when such events are generated from a large number of independent sources. A typical Poisson random variable is a count of the number of events that occur in a certain time interval or spatial area, e.g. the number of calls received by a switchboard during a given time period.

For the Poisson distribution to be applicable, several conditions must apply:

The events of interest occur at random over a particular continuous period of time (or distance interval, region of area, etc).

Events occur singly, i.e. not exactly simultaneously.

Events are statistically independent, i.e. the occurrence or non-occurrence of an event does not affect the chance of another event occurring. Hence, the occurrence of events is 'memoryless'.

Events occur at a constant average rate, usually denoted by the Greek letter λ.

With these conditions, it can be shown that:

Probability (r events in time t) = $(\lambda t)^r \, e^{-\lambda t} / r!$ r = 0, 1, 2...

The Poisson distribution has expected value λt and an identical variance λt.

The statistical confidence intervals – again, found in standard textbooks – for a Poisson distribution are of key importance in assessing data on rare events. A key calculation is the one-sided upper confidence point at 95% – the value "a" of λt for which the observation of 0 events is 5% probable, ie:

$$\text{Probability (0 events in time t)} = a^0 e^{-a} / 0! = e^{-a} = 0.05,$$

which gives a = 2.995… – say 3.

So, if during a period, zero events are observed, then with 95% confidence it can be said that the mean value for observation is less than 3. This is a powerful message: the absence of events provides 'weak' evidence about the underlying rate.

In the present context, suppose one wanted to confirm that the value of λ according to the TLS, 1.5×10^{-8}, was being achieved in practice by counting the rate at which accidents occur. Note that these are collisions (and flying hours) in all the en route airspace under consideration (see Davies and Sharpe (1993) for definitions), and that in this simple model it is assumed that about the same degree of protection against collision risk is to be assured everywhere in this airspace. Observing for 1.33×10^8 flying hours (a large number – 1.33 times 100 million) would statistically be expected to produce two accidents, i.e. one mid-air collision – noting that it is collisions which are Poisson events, not accidents. However, an absence of collisions would be the most statistically likely scenario, which would imply an upper confidence level of 3 mid-air collisions in 1.33×10^8 hours, i.e. a collision rate of 2.25×10^{-8}. Thus, even observing for long periods comparable with the expected interval between accidents does not produce strong confidence in the value of the ALS. Observations for much shorter periods are far worse: no mid-airs in a period of the order of million hours produces an upper confidence level two orders of magnitude above the TLS.

Rare events pose an intrinsically difficult problem for the system designer. Even if one or two accidents were to be observed, this does not produce large improvements in the degree by which he or she could be statistically confident about the TLS being achieved. The upper confidence levels for 1 and two observed events are shown below.

Number of observed events	**Upper Confidence limit at 95%**
0	3.00
1	4.74
2	6.30

These results, obvious to a statistician, show the impossibility of statistical proof through direct observations on a new system, under controlled experimental conditions, that ASAS meets the TLS. The use of methods other than classical statistical inference does not help. Subjective probability techniques (Hacking,

1976) depend on an individual's understanding of relative outcomes, and are hence inappropriate for rare events. Bayesian probability techniques require some kind of prior probability distribution to be constructed (the conjugate prior distribution for a Poisson is a Gamma distribution) – but where does the prior knowledge (*sic*) of rare events come from?

There is even no way of proving the safety of the existing ATM system statistically (indeed, the 2002 mid-air collision over the Swiss-German border could be said to indicate that the TLS is not currently being met in European airspace). There are obvious problems for 'verification', defined earlier.

To solve this problem, system safety somehow has to be partitioned into elements, for each of which safety factors can be quantitatively demonstrated, with the necessary statistical confidence, for the RES. Ideally, each element in this partition model would provide safety factors of 10^2 or 10^3, so that their product would deliver the required TLS. Thus, the risk has somehow to be built up of separable components, i.e. be the product of statistically independent events or characteristics. Two modelling methods can be envisaged: either the model is built up from features of the present system plus reliably modelled behaviour for new features or the model is constructed from the characteristics of the future system using some kinds of general principles. These may be termed Collision Risk Modelling (CRM) and Probabilistic Risk Assessment (PRA), although there are wide variations in the research literature. To these must be added real-time simulation of different components of new ATM systems. [FAA/Eurocontrol, 1998 is an excellent review article on these and other aspects of collision risk estimation.]

These kinds of issues are not specific to aviation. The nuclear power plant industry has had an ongoing debate over much of the last half century about probabilistic risk analyses. A quote from Yellman and Murray (1995 – see also Watson, 1994 and 1995) makes the point vigorously and succinctly:

> "Now consider the statement 'I estimate the probability of a core-melt in this (only partially designed and as yet unbuilt) nuclear power plant over ten years of operation to be 0.0000346'. Not only does the plant not exist, it may never be built, or at least not built to the current design. And if it is built,…we won't get a large and stable enough sample of its operation to validate the 'estimate' to any meaningful degree of confidence. It sounds less abrasive to say that a probability 'estimate' is being made than that an assertion is being made."

Therefore, the message is that ASAS 'risk estimates' are actually assertions about the degree of risk. There is no way of demonstrating by logical or statistical means that the TLS 'standard' would be met – the best evidence that can put before the RES is a well argued and robust assertion about the ALS.

4. SEPARATION MINIMA

One of the key safety barriers used to protect against mid-air collision is the use of separation minima (sometimes referred to as separation standards). Any sensible theory or framework for mid-air collision risk has to provide an understanding of the steps required to get from considerations of separation minima to estimates of that risk. Background on separation minima is given in Brooker (2002) and FAA/Eurocontrol (2001). Their role in ATC is open to several interpretations (to quote Simpson (1998): "ATC separation criteria is usually an area of confused and non-rigorous analysis"). From a system point of view, separation minima are 'formal rules'. Originally, these standards were required because of inaccuracies in radar and altimetry data, but they are increasingly seen as 'buffers' to permit effective warnings and controller/pilot actions. For present purposes, they are the minimum distances that controllers should permit between aircraft.

For example, in airspace with secondary radar coverage, the controller operates with 5 Nm plan (= horizontal) separation and 1000 feet vertical separation; at least one of these minima must be being achieved at all times. In the case of vertical separation, the minimum is interpreted as being at least one flight level apart. There are airspace regions that are not covered by ground-based ATC – the North Atlantic (NAT) Region is an important example. The NAT region operates using a structured track system with (eg) aircraft on adjacent tracks at the same flight level are kept 1° (roughly 60 Nm) apart. This standard was determined by consideration of navigation performance rather than 'active' ATC - the driving force to reduce separation minima is fuel penalties. ASAS concepts used in the NAT region could help to reduce separation minima further, but this is not discussed in the following.

For ASAS in en route airspace, to quote the FAA/Eurocontrol (2001) document

> "…it is supposed that airborne separation will be provided and maintained by flight crews applying standardised separation minima. Therefore, the major issue is the establishment of these 'airborne separation minima' so as to achieve safe flight operations. Optimistic views are that airborne separation minima could be much smaller than ATC radar separation minima and could thus allow for capacity increases. Other views are much more reserved and warn that they might be larger than ATC radar separation minima, while possibly smaller than procedural separation minima…These separation minima will have to be established at the ICAO level..."

There are two extremely good reasons for retaining the existing minima in an ASAS system, certainly initially in the RES's work (aside from the complexities involved in trying to make changes, e.g. see the recent work on separation minima changes in an ASAS environment (Eurocontrol, 2002b)). The first is elegantly set out by Quine (1987 – see also Simon, 1982), the second by Popper (Medawar, 1991). Quine notes the virtue of constraints as a 'freedom from decision': it limits the searches required. At present, the controller does not have to carry out an assessment about the likelihood of a mid-air collision between pairs of aircraft: all he or she need do is to ensure that the separation minima are not breached. Thus, a decision about what to do – and hence the possibility of some kind of error – is replaced by an effective 'rule of thumb' that establishes 'protection criteria' between aircraft (Simpson, 1998).

Popper's point is that it is wise to retain as many features of the existing system in the new system as possible, and try to introduce change step by step, thereby minimising the problems and complexities introduced by new interactions in human-machine systems: 'piecemeal social engineering' is to be preferred to full scale re-engineering. In particular, to change the arrangements for separation minima would risk burdening pilots with extra tasks beyond those that ASAS would 'transfer from the controller'.

Another viewpoint, focusing on en route capacity, is that the economic gains from ASAS could be very substantial in terms of reductions in ground-based operational costs and the potential for increased capacity (see, for example, Brooker, 2002a and 2003). These would far exceed any from improved flightpaths arising from presumably marginal changes to the separation minima. So why 'tinker' with the latter and hence put in jeopardy the early introduction of ASAS? Strategic financial decisions about ATM system-level investments are inherently difficult ones, because the operational adaptive skills of ATC and airlines tend to reduce or defer new technology benefits, while few major changes produce immediate operational cost savings and revenues to airlines and airports (see Brooker, 2002c).

These are different, perhaps somewhat abstract, ways of viewing the issue, but the practical safety benefits of retaining separation minima can perhaps be simply illustrated from a probabilistic viewpoint.

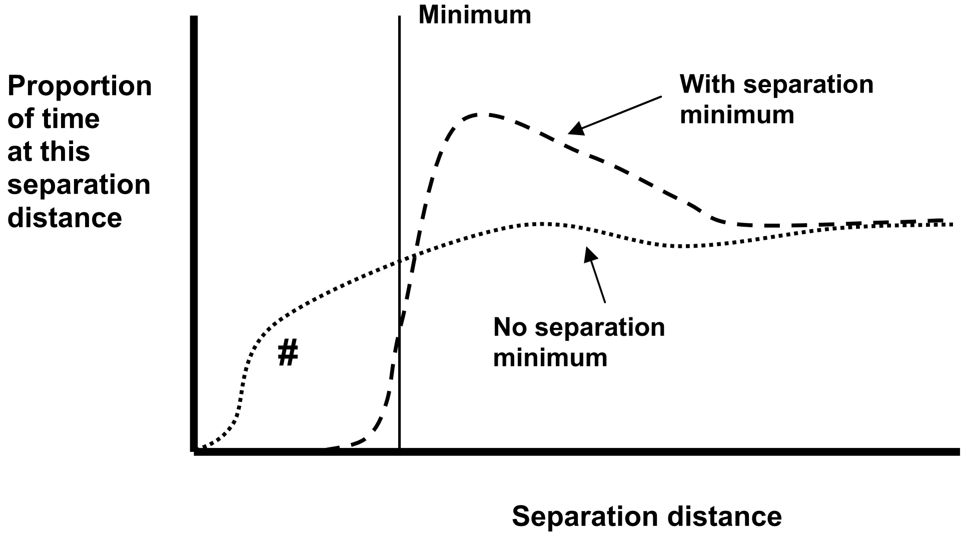

Figure 1. Effect of separation minimum of distance distributions – speculative

Figure 1 speculates – note the word – on how the use of a separation minimum changes the proportion of time potentially spent at close distances. The dotted line is the case without a minimum, with ATC intervening only when controllers judge it is 'necessary' to prevent a potential serious conflict. Controllers would probably work with some minimum miss distance in mind. The closest separations would occur when they misjudged relative velocities or when there was insufficient time to instruct pilots to manoeuvre. There would be quite a few moderately low separation values (at #) – circumstances where the aircraft were close in position but in which the configuration, given relative velocities, would not be judged hazardous. These could not be allowed if a separation minimum were in operation. The dashed line shows the distribution of distances between aircraft when using a minimum. The effect of the minimum is to move the aircraft pairs at 'near to zero' separation to 'greater than or near to the minimum'. Thus, the key point is that risk calculations now focus on deviations from a safe value rather than the closeness to an unsafe one. This offers an extra probabilistic layer of safety. Such a role for separation minima has probably been recognised in the past, but its importance has not received much attention.

5. PROBABILISTIC RISK ASSESSMENT

There is a huge literature on Probabilistic Risk Assessment – PRA (Probabilistic Safety Analysis – PSA – is also used). PRAs try to estimate the risk of accidents by analysing the sequences of events that could produce an accident – the 'causal chains'. Failures arise from 'errors' – or indeed natural variations – from 'normal operations'. At each stage, the probability of an event's success or failure in safety terms has to be quantified. For events representing the function of mechanical or electronic components the failure probability can in theory, and often in practice, be determined by observations of the performance of that particular sub-system. An understanding of failure modes and engineering characteristics leads to a valid PRA estimate. But complex systems generally contain people who have make decisions and act on the information presented to them, so some of the events require probabilities to be estimated for 'human components' – the task of Human Reliability Analysis (HRA).

While these kinds of probability might well be calculated to a good degree of approximation for an industrial process, it is much more difficult to produce estimates of these risk components for what are already rare events. A necessary ingredient is that the mechanisms and factors involved should be traceable to what happens in the real world. The models may well be appropriate but the difficulty is in 'populating' them with relevant data. An example is Foot (1994), a trial PSA of the present (*sic*) UK CAA en route air traffic operations. This study illustrates how classical hazard analysis techniques might be applied to subsystems of a complex, man-in-the-loop system to obtain the collision risk. It also demonstrates the combinatorial explosion in fault tree complexity and hence the requirement to estimate a raft of failure mode parameters.

The traditional way of getting around the problem of the inherent uncertainty in probabilistic risk assessment is to aim for a cautious assessment. If it is possible to show that safety targets would be met, even when ignoring significant safety barriers (such as TCAS – of which more later) and overestimating failure rates, then the problem goes away. Unfortunately, this seldom works even with current ATM systems: a collection of 'cautious' assumptions generally tends to produce over-pessimistic risk estimates, and hence has little value for safety decision-makers.

These difficulties with HRA methods, particularly in the nuclear power plant case, have themselves generated a huge literature. Much of the impetus for a very critical approach to the subject came from a special edition of 'Reliability Engineering and System Safety' in 1990. The editorial by Dougherty and the papers by Swain and Moray are good examples of the honest – and indeed wise – analysis of the issues. Hollnagel's book (1993) and the more recent NATO conference (2001) are two instances of the continuing debate on the topic.

Dougherty set out the problems with HRA simply:

> Insufficient empirical data
>
> Concerns about use of expert judgements, particularly for rare events
>
> Lack of confidence that simulator data matches real life
>
> Disconnects between modelling assumptions and psychological knowledge
>
> Use of 'Performance Shaping factors' to modify data

Indeed, some behavioural scientists tend to believe that a much deeper theoretical foundation is needed before quantification should be attempted. This is a perfectly valid view, but not very helpful to system designers – how long do they (and society) have to wait before the imprimatur of the researchers? Hollnagel (1993) does not discuss safety targets, but comments, re improved technology to improve ATC capacity:

> …all these enabling technologies could have been developed and used to reduce the level of risk while keeping system utilisation constant. In other words, although flights would not have become more frequent, it would have become safer to fly."

Again, this is a position that can be rationally held, given appropriate premises, but is it fruitful for airlines and ATM system designers? The public actually wants better safety and more flights– the TLS has been progressively reduced over the last 30 years.

But HRA's issues and lessons cannot be ignored. Some quotes from Moray (1990) are very relevant to the ASAS case:

> "The use of 'expert judgement' is a polite name for 'expert guesses', and we do not have data to validate the accuracy of the guesses.
>
> The attempt to find a single number is an attempt to establish a context-free universal fact about human performance. No such thing exists. It is simply fantasy to think that the probability of human error is described by a single number…
>
> [Re flight crew error rates] How do any of us survive? The answer (and the lesson) is that in the case of airliners they are quite forgiving systems…there is time enough, usually, to make errors, discover them, and recover from them.
>
> The most serious design deficiency of the Chernobyl reactor was that…the instability was such that once it occurred there was no time to recover from the error."

A PSA incorporating a HRA is thus a complex – and ultimately probably correct model – that is likely only to produce usable answers at some indefinite point in the future. This cannot be the way forward in the present circumstances, and 'giving up' is surely not the conclusion that should be drawn. There is obviously a great deal of valuable information derived from Human Factor experiments. How can this best be used and developed? Can a simpler model framework be constructed that delivers testable results?

The danger is therefore the creation of 'over-elaborate' models – ones whose parameters cannot be reliably estimated from the data likely to be obtainable. Thus, the focus has to be on the simplest model that can be soundly based on available data. Section 8 presents a possible model, but it is first necessary to explore some human factors-related aspects of Airproxes and ASAS.

6. 'REASONABLE INTENT' RISKS

Brooker (2002) discusses the lessons that can be learned about potential collisions by studying Airproxes; these lessons from Airproxes and other safety incidents are not sufficient to prevent all future types of accident – but they do offer some necessary tests. The UK Airprox Board (UKAB) Report data for the year 2000 revealed:

(a) No incidents arose because of a radar accuracy or resolution problem, or from normal altimetry operation. This is not to say that these do not exist, but, on the statistical evidence here, they would be a causal factor in only a small proportion of Airproxes. Scans of earlier years' Airprox reports reveal a similar picture.

(b) In no case did pilots or controllers express strong concerns about disruption by warning systems.

(c) Most of the Airproxes were judged by the UKAB to have been caused by failures in procedures, rules, structures, or communication by both pilots and controllers, e.g. 'pilot misunderstood the ATC instruction'; 'ATC and pilot procedure errors'; 'controller distraction when sector split'; 'pilots' poor RT discipline'; 'pilot procedure for altimetry in error'; 'ATC memory slip'.

(d) Past UKAB Reports list as the top four causal factors: 'did not separate/poor judgement by controllers', level busts, 'did not pass/late passing traffic information', poor coordination by controllers. These immediate causes in this UKAB categorisation would all seem to have been human errors of some kind.

Brooker (2002) then introduces two concepts relevant to mid-air collision: 'Position Integrity' and 'Reasonable Intent':

> Position Integrity: The system has this when positional equipment for navigation and surveillance is functioning 'normally' – when the errors on radar, GPS, altimetry, measurements are not extreme, when displays work properly, when signals are not corrupted or lost, etc.
>
> Reasonable Intent: this is an inference that would usually be made 'after the event': did the controller implement what a competent controller would have considered a reasonable (albeit perhaps not perfect) course of action; did the pilot do something that other pilots would have judged decent practice (albeit perhaps not the ideal decisions)? It thus covers misjudgements and blunders as normally understood. It is primarily a human factors issue.

In risk budgeting terms (i.e. to compare with the TLS), if Airproxes are a good guide to potential mid-airs, risks from Position Integrity failure are the kinds of event resulting from the loss of planned separation for which 'traditional' (See FAA/Eurocontrol, 1998 for a historical review) collision risk models were developed. These models largely focused on equipment performance or failure, which would be 'attributable to the loss of correctly established separation'. In practice, these collision risk models had to be adapted to incorporate reasonable intent errors. For example, even 20+ years ago, the NAT model North Atlantic had to include waypoint insertion errors as well as navigation errors (Brooker and White, 1979).

Reasonable Intent failures are more typical of observed Airproxes, generated by human error in the widest sense. In crude terms, the first is 'wrong place on right flight path' and the second is 'right place on wrong flight path'. [These types of events are both 'first order', in that they reflect what are essentially single failures, but there can also be multiple problems and specific emergencies – see Brooker (2002) - well covered by safety regulation – as evidenced in Profit (1995), and are therefore not discussed further here.]

The observed relative proportions of different types of Airprox suggest that, for the current system, comparatively large proportions of the Air Traffic Services risk budget should be allocated to Reasonable Intent risk. The key argument here is that mid-air collision in an ASAS environment will predominantly arise from the latter type of risk.

These types of risk cannot be calculated by a similar means to the Position Integrity risks. With these 'equipment' risks, it is possible to analyse relevant data of what are essentially engineering observations. By their very nature, Reasonable Intent risks are deeply embedded in human functions and performance – the subject of Human Reliability Analysis (HRA). The problems of HRA's application to ATM therefore need to be solved. This might well be a formidable undertaking: controllers' tasks appear to be much less structured – i.e. with more discretion to determine solutions – than are nuclear workers'.

The extent to which the risk in an ASAS environment would reflect the current situation – where reasonable intent failures dominate – depends of course very much on how the ASAS system is designed and what is the precise role of the human. At one extreme, in particular in an 'electronic Visual Flight Rules' environment, it is likely that human errors would be the dominant concern. However, if the ASAS equipment offered solutions to conflicts, or even executed some types of these solutions automatically, then the equipment performance would obviously be a much more important factor. The Airproxes examined occurred in a human centred operation, where equipment provides information for decision-making but seldom offers solutions and never 'makes the decisions'; hence it is hardly surprising that human errors were leading causal factors.

7. ASAS SAFETY ISSUES

Why should one believe that an ASAS operational concept could be constructed that would at least maintain and preferably increase safety? The following derived from FAA/Eurocontrol (2001) notes some key safety points:

Situational Awareness: The flight crew is presented with flight information concerning surrounding traffic, possibly in conjunction with a navigation display or a surface map, which assists flight crews with:

- see-and-avoid duties;
- avoiding blunders or errors; and
- information to facilitate correct decision-making.

Automation: ASAS uses various sources of position and intent data. ASAS generates guidance to the crew for safe and timely resolution of conflicts or maintenance of safe separation.

Guidance presented directly to flight crew: ASAS guidance does not depend on ground-to-air communication, hence preventing the common hazard of missed or garbled radio communications.

But there are potential negative effects on safety. The controller's tasks are either eliminated or transferred to the pilot. Could significant features be lost? Could there be new types of pilot-induced errors, or increased rates of existing errors, perhaps due to increased workload?

Human Failure Modes

A great deal of work has been undertaken in recent years by Eurocontrol and states to develop tools and methodologies for the analysis of human error in ATM incidents focusing particularly on Human Factors in the resolution of incidents. Key elements are the identification of the forms of human error that occur as part of an incident, and the decomposition of these errors to determine the psychological mechanisms behind the error, and hence the reasons why the errors occur. One recent technique is TRACEr (Shorrock and Kirwan, 2002): TRACEr provides some basic 'grammar' for understanding errors.

At a high level, errors fall into a number of categories associated with the task that is being performed (e.g. radar monitoring, strip handling, etc). Each of these errors can have a number of underlying causes (e.g. judgement, planning/decision-making failure, perception and vigilance failures). The ultimate cause of an error is the psychological mechanism that results in the operator making an error. Such mechanisms include perceptual tunnelling (when the operator focuses on one particular situation at the expense of all others) and information processing failure (where the operator's information processing system is unable to cope with the type or quantity of information presented). For the purposes of the analysis of human errors in ATM incidents, a taxonomy has been developed for task errors (Shorrock and Kirwan, 2002), shown in Figure 2 below.

PRESENT ATM SYSTEM

Controller Task Errors

Separation
Controller-pilot communications
Radar monitoring
Aircraft observation / recognition
Co-ordination
Control room communications
Aircraft transfer
Flight progress strip use
Operational materials checking
Training, supervision, or examining
HMI
Other task
Pilot tasks
Pilot-controller communications
Aircraft handling
Visual observation
Flightdeck co-ordination/communications
Operational materials checking
Training, supervision, or examining
HMI
Other task

ASAS SYSTEM

Pilot Task Errors?

Pilot-controller communications
Aircraft handling
Visual observation
Flightdeck co-ordination/communications
Operational materials checking
Training, supervision, or examining
HMI
Separation
Radar monitoring
Operational materials checking
Other task
NEW – ASAS-RELATED

Figure 2. Task Error Taxonomy

‘Separation Error’ needs explanation. These are errors associated with controllers climbing, descending or turning aircraft into conflict with other traffic (compare with ‘Reasonable Intent’ risks above). ‘Separation Error’ does not therefore refer to the outcome of the error, but rather an error in which the prescribed separation was not maintained and which was not detected in time to prevent the loss of separation.

What extra types of error would be produced in an ASAS environment? Would there be increased rates for existing types? Could pilot workload be a significant issue? How loosely or tightly coupled should the system be? Taking a key thought from Moray (1990): how ‘forgiving’ would the system be of errors and failures?

8. A 'MINIMAL FRAMEWORK' COLLISION RISK MODEL

Definition of the Model

To try to understand how an ASAS safety proof might be developed, it is easiest to construct a 'Minimal Framework' Collision Risk model. This probabilistic framework establishes logical connections and allows the several distinct components of modelling to be examined. It starts from aspects clearly related to breaches of separation and then focuses down to collision risk. The development of the framework makes apparent what are the key Human Factors experiments that need to be performed, i.e. it powerfully links in these experiments to risk estimation. First, it is necessary to define some terms, which are then analysed.

A 'Significant Separation Breach' – SSB - is an event when two aircraft have lost separation in all dimensions by a significant amount and action may be required to recover minimum separation. 'Significant' means that small deviations from minimum separation, e.g. because of wind effects or FMS smoothing, are tolerated: for the moment. However, 'significant' does imply that the breach is unacceptable – but not necessarily 'unsafe' or 'hazardous'. If the present system is any guide, SSBs will very probably occur through blunders and other kinds of Reasonable Intent failure rather than Position Integrity problems (although it will still be necessary to estimate the risks from the latter type of failure – e.g. Brooker (2004a/b) discusses analytical calculations of collision risks arising solely from radar inaccuracy). These 'proximate' incidents are prima facie instances of some degree of need for action to resolve possible problems – potential 'precursor' events. Simpson (1998) offers some interesting examples of potential encounter criteria akin to an SSB. As defined, an SSB would roughly correspond to Simpson's 'Monitoring Criteria'. An SSB might also be compared to an Airprox (see Brooker, 2002 for references and discussion). The Significant Separation Breach Rate – 'SSBR' – is the frequency of SSBs per number of system flying hours.

The next definition is the SSB/Collision Scaling Factor – SF for short. It is the ratio of the long-term average of collisions to SSBs <u>if</u> no safety defensive barriers were in place, i.e. 'blind flying'. It reflects the fact that an SSB has much larger dimensions than an aircraft, and so the SF is much less than unity.

The third definition tries to summarise detections and actions in the 'defensive barrier' process: the 'Barrier Failure Probability' – BFP. Two families of probabilities need to be estimated: P(T), the probability of detection by time T before closest approach – through ASAS, TCAS and visual acquisition – and Q(T), the probability of effective conflict resolution given an alert at time T.

With these definitions of terms, the rate of collision CR is:

$$CR = SSBR \times SF \times BFP, \text{ ie}$$

$$CR = SSBR \times SF \times \{ P(T) \# (1 - Q(T)) \}$$

Here the hash sign and curly brackets represent the convolution integral of the functions P(T) and (1 – Q(T)), i.e. the probability densities are multiplied and the summed. The picture is a simple one – Figure 3.

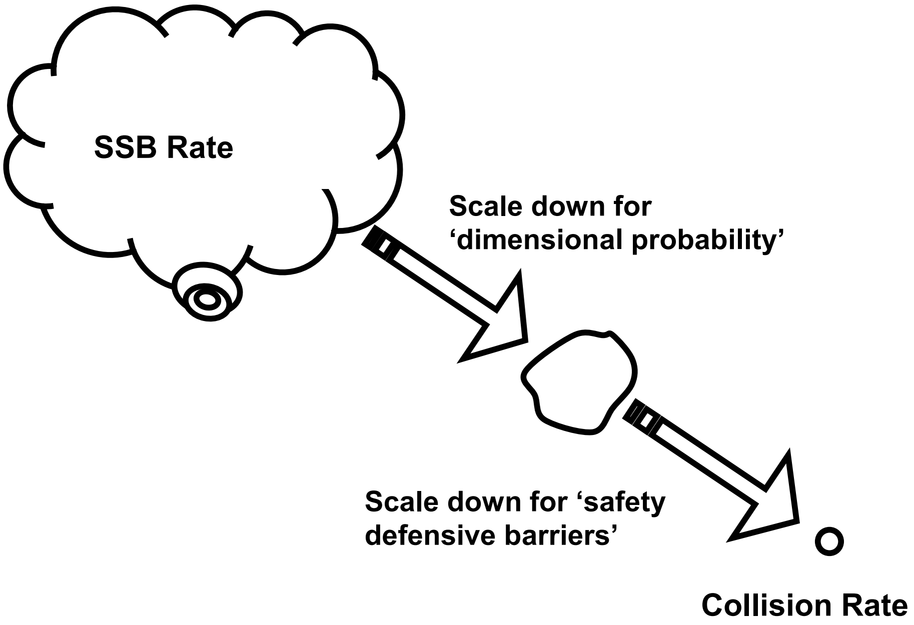

Figure 3. 'Minimal Framework' collision risk model illustration

This model does need some refinement before it is a full description. For example, there is a possibility that an 'unnecessary' evasive manoeuvre – i.e. where the aircraft would not have collided if they had kept to their previous flightpaths, could result in a collision. This is probably relatively unlikely, but would need to be handled through a separate analysis. Relevant research has been carried out by Geisinger in the context of the Analytic Blunder Model (referenced in FAA/Eurocontrol, 1998). Also relevant is the work done to estimate the rate of 'induced collisions' that occur through the use of TCAS (Harrison, 1993), which showed that in ideal circumstances these could be expected at a rate of 4% of the existing (*sic* – i.e. before TCAS) risk of collision.

Estimating the Minimal Framework Model parameters

Significant Separation Breach Rate – SSBR

The en route horizontal and vertical separation minima are taken as 5 Nm and 1000 feet respectively. These tend to the values used by ATC providers with (eg) monopulse secondary radars. For an SSB to be defined, there needs to be some judgement about what constitutes a 'significant' breach of these minima. As noted above, Simpson (1998) has examined this kind of issue with some rigour. For present purposes, the assumption is made that a horizontal breach of 1 Nm and a vertical breach of 300 feet would count as ‘significant’. These figures are ‘not unreasonable’ given the levels of navigational performance currently being achieved.

How would the SSBR be estimated? As already noted, SSBs very probably will not occur through Position Integrity problems but from failures of Reasonable Intent – from the consequences of human error modes of the types discussed above in relation to TRACEr. The Sherali et al (2000) and Barnett (2002) modelling work are starting points. They focus on the rate at which conflict probes will be required to help resolve potential conflicts. The next stage used to use a technique such as TRACEr (Shorrock and Kirwan, 2002) to help to generate realistic and comprehensive SSBs at reasonable rates. TRACEr provides some basic ‘grammar’ for understanding errors but further modelling work is required. These are easy to write but the research tasks are probably very much harder to do in practice – but it has to be done.

Unfortunately, Airprox data is unlikely to be an adequate guide for SSBs. The present ATC processes generate Airproxes from what is largely a fixed route system plus tactical procedures, e.g. with pilot/controller negotiated climbs and descents. ASAS airspace and procedures would be very different.

The range of airspaces used in this modelling would need to cover the full range of types of future airspace volumes and traffic densities/patterns. Given that traffic is expected to increase considerably, ASAS would need to be able to demonstrate it can cope with high frequencies of potential conflicts in highly dense airspace. It would be essential to include factors that require significant ‘non-direct’ routeing. An example is a military zone, either permanent or temporary, which in some States these can occupy sizable airspace volumes). Weather-related restrictions on particular routes and locations would also be ‘natural features’ in worldwide ASAS airspace.

The TLS is derived for flights using a large volume of en route airspace. The risks in particular parts of that airspace can be different from this ‘average figure’ – see Brooker and Ingham (1977). However, large deviations from the average TLS would not be tolerated. In ASAS safety calculations, the risks in the different types of airspace – in the broadest sense – would therefore need to be properly weighted by their system flying hours. It cannot be assumed that SSBs would always occur ‘randomly’ throughout the airspace concerned. For example, SSBs might be more likely near a particular boundary point in ASAS airspace, perhaps where aircraft would be ‘funnelled’ into traditional airspace. This could generate incidents where aircraft were cleared to already occupied flight levels, i.e. the aircraft would level off at the worst possible location in risk terms (but this higher risk would tend to occur for a comparatively small proportion of flying hours).

SSB/Collision Scaling Factor - SF

SF is defined as the ratio of the long-term average of collisions to SSBs if no safety defensive barriers were in place, i.e. 'blind flying'. It reflects the much larger dimensions of an SSB compared to an aircraft.

Suppose aircraft SSB dimensions are modelled as discs (see Brooker (2002 for more detail on this model and references to the earlier original work by May and others). Each SSB disc has height H_b and radius R_b as shown below.

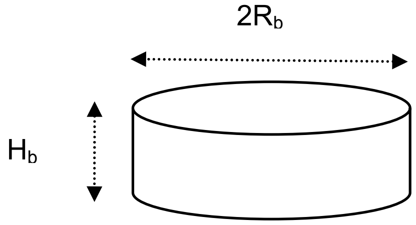

Figure 4. SSB disc

Aircraft are taken as having an SSB if their discs touch. Next, assume that discs are always orientated along the relative velocity vector $\mathbf{V_r}$ (magnitude V_r) of the two aircraft – given that aircraft generally do not change altitude at extremely high rates of climb. [In many collision risk models, these discs are orientated in the normal 'xyz' coordinates, but this is just a convention. The discs are not 'real' – they just serve to 'envelop' the aircraft – so their orientation in the model is a matter of choice. The choice made here is convenient and has no physical or safety significance, i.e. it does not influence the validity of the results.] Now, model an SSB by an equivalent picture – taking aircraft 2 as a point and the 'protected zone' of aircraft 1 as a larger 'collision disc' of dimensions $2H_b$ and $2R_b$, in the frame of reference based on aircraft 2, i.e. in which it is at rest, as illustrated below

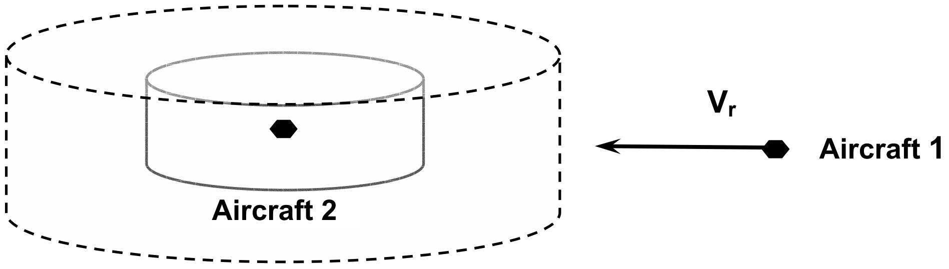

Figure 5. Equivalent collision disc

The 'cross section' of the larger collision disc viewed from Aircraft 1 is simply a rectangle of area proportional to H_b x $R_{b.}$ If the positions and velocities of aircraft in an SSB are statistically random, then the calculation of SF is one of geometrical probability. The dimensional dependence for collisions/SSBs would be in proportion to H x R, where H is the height of an aircraft and R the radius of a collision disc just enclosing its fuselage and wingspan (Brooker, 2002 and May, 1971).

Taking the SSB dimensions as:

H_b = 700 feet = 0.115 Nm

R_b = 4 Nm

and the aircraft disc as:

H = 0.010 Nm = about 60 feet

R = 0.017 Nm = about 100 feet

gives:

SF = (H x R) / (H_b x R_b)

= 0.010 x 0.017 / 4 x 0.115 = 1 / 2,705

But this very small number is not a realistic estimate, in particular because of the high accuracy of height keeping.

Currently, thanks to improvements in altimetry, aircraft in cruise tend to fly very close to their flight levels (see Brooker (2002) and Moek et al (1993) and use only a few of them (see Brooker (2002), quoting Moek et al, 1993). Most modern jets are optimised to fly at around the tropopause, say 35,000 feet, so FL 350 and the immediate neighbouring flight levels are very popular. The data from Moek et al (1993) on vertical errors is: for long/medium range types, the standard deviation was 85 feet; for medium/short range types it was 155 feet. Since 1993, there have been continued strenuous efforts by airlines and ATC providers to improve and monitor vertical performance, as part of the programme of work to reduce vertical separation above FL 290 to 1000 feet.

However, in a free flight ASAS scenario using existing separation minima (e.g. see Hoekstra et al, 2002), acceptable flightpaths for aircraft on crossing routeings would be separated by one or more flight levels. This vertical concentration means that the likelihood of the aircraft being in 'vertical overlap' would be much the same for both SSBs and collisions, i.e. there would be much less 'dimensional scaling down'. A simplified calculation of 'vertical overlap for aircraft flying at the same level shows the importance of the effect and the nature of the functional dependence.

Denote the aircraft height by H and the probability distribution of heights about the flight level by $f(v)$. For a second aircraft at a distance Y from the flight level the probability of the two aircraft being in vertical overlap is:

$$\int_{Y-H}^{Y+H} f(v)dv$$

If this is summed over all possible Y values, this gives the probability of vertical overlap as:

$$P_z(0) = \int_{-\infty}^{\infty} f(Y) \int_{Y-H}^{Y+H} f(v)\, dv\, dY$$

The zero in the function is just a reminder that the aircraft are intended to be at the same altitude z. It can be assumed that f is a well-behaved function that can be expanded out in a Taylor series, to give:

$$P_z(0) = \int_{-\infty}^{\infty} f(Y)\, \{f(Y).2H + cubic\ terms\}\, dY$$

The quadratic terms cancel out in the expansion. If the cubic terms can be neglected this becomes:

$$P_z(0) \cong 2H \int_{-\infty}^{\infty} [f(Y)]^2\, dY$$

The integrand's dependence on $[f(Y)]^2$ is important. As an illustration, taking $f(Y)$ as double exponential (not too far from Moek et al's observed height keeping data) gives:

$f(Y) = e^{-|Y|/\lambda}.[2\lambda]^{-1}$ [NB: The standard deviation of $f(Y)$ is $\sqrt{2}\lambda$.]

This gives:

$$P_z(0) = H / 2\lambda$$

Thus, an aircraft disc of height 60 feet and a λ value of 60 feet would give a $P_z(0)$ of 0.5. In practice, $f(Y)$ could be estimated in the same ways used in earlier studies (eg, Moek et al, 1993), so the integrals could be computed numerically – and might well show that the cubic terms above should not be neglected.

In the horizontal dimension, the ratio of R_b to R would be 4 to 0.017, i.e. 235. Multiplying this by 2, i.e. assuming independence between vertical and horizontal probability distributions, gives SF as about 1 in 470.

Note the need for 'randomness' in this calculation. It is possible to think of ways in which the relative velocity vector might not be uniformly distributed in the R_b dimension, e.g. with the funnelling effect noted earlier.

Barrier Failure Probability'

The previous elements in the calculation presented – 'blind flying' in the fullest sense – do not include the safety defensive barrier effects of automatic warning systems and controller/pilot action, including the effectiveness of See-and-Avoid. BFP measures the effectiveness – or rather the possibility of failure – of these barriers. The present system has three main safety barriers:

Conflict Alert (STCA) – The ground computer processing system has the facility for analysing SSR tracks to predict if aircraft might come into close proximity in the near future and, if they do, warn the controller by flashing a message on his radar screen. Subsequent controller instructions would be ‘normal’, e.g. would generally tend to separate the aircraft by horizontal vectoring.

Traffic alert and Collision Avoidance System TCAS (generically ACAS II) [The abbreviation TCAS will be used here, except in quoted text, to emphasize the difference between ASAS and TCAS] – An on board collision avoidance system based on detection of other aircraft in the vicinity carrying SSR transponders. These tell the pilot of nearby traffic – TA (Traffic Advisory) – and aircraft coming into conflict – RA (Resolution Advisory). RAs tell the pilot to climb or descend as appropriate to take it out of risk with immediate action.

‘See-and Avoid’ – The pilot visually searches for other aircraft, and then changes course if this is necessary to avoid them. Aircraft crew are exhorted to maintain vigilance so as to see and avoid potential mid-air collisions.

Other elements also contribute to safety of course, such as the `party line' effect in voice communications and the crew's experience about what generally happens at particular locations and routeings.

With Full Delegation, ASAS takes over the role of STCA. ASAS is used to ensure that the separation minima between aircraft is maintained; TCAS provides a final independent backup system if the separation minima are breached and there is risk of collision. See-and-avoid would be the ‘last resort’ safety tool. ASAS conflict detection would operate 10 to 20 minutes before the closest point of approach (CPA), while TCAS provides traffic advisories 20 to 50 seconds before CPA and Resolution Alert (RA) gives warnings 35 seconds before CPA (Eurocontrol, 2002a). ASAS would need to be integrated – both in terms of equipment and in operational usage – with an aircraft’s TCAS (e.g. see Abeloos et al (2000), Zeitlin and Bonnemaison (2000), and Zeitlin (2001), and later text in this section regarding independence and common mode failures).

Should the effects of al these three be taken into in the risk calculations to estimate an ALS to be compared with the TLS? These are examined in reverse order.

A major study on See-and-Avoid (BASI, 1991) concluded that in visual conditions, in the absence of some form of traffic alert, the probability of a pilot visually acquiring a threat aircraft is generally low until a short time before CPA. For commercial aircraft speeds, See-and-Avoid usually failed to alert potential collisions. Even under the best conditions, visual search can be like ‘looking for a needle in a haystack’, and in poor visibility the chance of it succeeding would be negligible. Thus, the pilot cannot reliably visually acquire other traffic or consistently. Trials under test flight conditions suggest that visual acquisition alone (i.e. without any ‘cues’ from an alerting system), is less than 50% effective (Moore, 1998).

Hence, See-and-Avoid, whilst it plays a useful role, does not reduce risk by a quantitatively significant amount. Moreover, for aircraft flying under IFR, it seems rather dubious to be reliant on non-instrument means for any part of the protection against catastrophic system failures – which a mid-air collision would certainly represent. Thus, there is not a strong argument for risk reduction from See-and-Avoid being estimated in the ALS.

TCAS is rather different – but there are some differing views about its ‘safety system function’ (FAA/Eurocontrol 2001). TCAS was introduced in order to reduce the risk of mid-air collisions; and has been designed to operate in all airspace. Thus – paragraph 2.2.6 – it presently ‘serves as a last resort safety net, irrespective of any separation standards’ appropriate to airspace categories, and ’it has no other role in the ATM system’. FAA/Eurocontrol (2001) notes that ICAO documents indicate that:

> The provision of ATC services in a given airspace shall not be based on the ACAS equipage of the aircraft; and
>
> Air traffic control units shall provide the same services to ACAS and non-ACAS aircraft.

On this basis, ATC procedures have to be judged safe without considering the effect of the TCAS safety net – so TCAS in practice helps to prevent mid-air collisions, but gets no ‘credit’ for this in system safety assessment.

FAA/Eurocontrol (2001) comments (paragraph 3.2):

> “Following appropriate clearances and instructions results in separation minima being maintained and thus ensuring safety. Separation minima are established such that the risk of collision is at an acceptable level. The other processes by which the flight crews avoid collisions also contribute to reducing the risk of collision, but they do so in an unquantified way…The use of ACAS does not amount to separation provision because it provides no guarantee that the risk of collision is reduced to an acceptable level.”

So, on this view:

> “In normal circumstances, when separation (ATC or flight deck) is provided, airborne collision avoidance should not be necessary. Applications implementing airborne separation should achieve the approved Target Level of Safety (TLS) independently from airborne collision avoidance.”

Although these views can be comprehended, they do not seem very rational ones in terms of the development of TLSs and collision risk modelling. TLSs and ALSs are by their very nature statistical statements rather than 'guarantees'. To reiterate an earlier point, collisions decades ago might have been more likely to be caused by equipment and navigation hardware problems, but today's Airproxes and other incident data shows that the highest likelihood for collision arises – in crude terms – from being in the 'right place but on the wrong flight path'. The view that separation minima somehow 'guarantee safety' by protecting against 'technical errors' on the flight path is therefore wrong. Separation minima of themselves do not guarantee safety, any more than a road speed limit prevents car crashes. It is actually the control of the 'failure rate' when minima are breached that delivers the required safety.

As already stressed, collisions are most likely to be caused by human error in the largest sense – and these would be very infrequent probabilistic events. Separation minima and TCAS alerts are different ways of providing a safety barrier against this possibility, and these barriers are, to different degrees, statistical in nature rather than providing 'guarantees'. One of them reduces the complexity of decisions that controllers have to take; the other alerts pilots and controllers to the need to take a decision. Both of them are now integral parts of the ATM safety system – so why should only one of them be included in risk calculations?

A major problem is that a flawed definition of 'ATM system' appears to have been adopted by Eurocontrol (e.g. in Eurocontrol SRC (2000a) – 'ESARR 4') and ICAO. Surely, the most rational definition would be something on the lines of:

> ATM system: Everything that contributes to the safe movement of air traffic – the 'Total System'.

The prime goal of the ATM system on this definition is to control risks. Safety in ATM is 'the interaction between Procedures, People and Equipment' (Baumgartner, 2003). The pilot is part of this ATM system. In the current system, the controller generally has greater knowledge of the ATM environment and the risks posed by neighbouring aircraft than does the pilot. But Total System safety depends on the pilot's actions, which include following instructions from the controller and that the pilot acts in accord with TCAS alerts.

Moreover, and essentially continuing the discussion in Section 3, the TLS was never intended to be a measure of 'acceptable air traffic control failure' but to be a target that the ATM system should achieve. The TLS was not developed on the basis that certain types of system, technology or procedure would either be present or absent. The risk calculations for an ATM system's ALS were seen as potentially including all mitigating factors, from controller monitoring and intervention to automatic warning systems. The TLS was not therefore produced in the context of the causal factors or mechanisms by which safety is either at risk from or by which it is assured. The ICAO teams that developed the TLS philosophy did not *a priori* rule out the use of systems such as TCAS in delivering the TLS (Brooker and Ingham (1977) sets out the key references). In the modern day, the point is well illustrated by Baumgartner's (2003) definition: "TLS: The level of safety which the total system is designed to meet".

These comments can only scratch the surface of the issues involved in setting the right future policy for the inclusion – or otherwise – of TCAS in hazard analyses. Further papers will endeavour to achieve a fuller analysis, with detailed critiques of present ICAO and Eurocontrol policy, including ESARR 4 (Eurocontrol SRC, 2000a).

Returning to the calculation of Barrier Failure Probability: the two generic terms P(T) and 1 – Q(T) in BFP, the first is, leaving aside See-and-Avoid elements, equipment-based, while the second depends on human performance given an alert. In both cases, these probabilities are averaged over the range of encounters. It needs to be stressed that all the parameters in the Minimal Framework Model are long-term averages for the airspaces under consideration – remembering that the TLS is itself a long-term average rate for mid-air collision accidents.

Estimates of P(T) can be made by simulating the operation of ASAS and TCAS on representative aircraft encounters. In the past, before these types of equipment were in common use, this could be done by using radar encounter data. A good example using UK data is the study by Hale and Law (1989). This was very important work because it showed inter alia that all genuinely 'serious' encounters, out of a sample of more than a thousand aircraft pairs, were detected by both systems. Its key conclusion was that 'the majority of conflicts likely to result in a TCAS RA would have been already alerted to the controller in good time to anticipate the RA'.

For the full delegation scenario, ASAS would have to be demonstrated to deliver at least equal performance to STCA, because it has to provide equivalent functionality. There are important issues here about the extent that ASAS and TCAS equipments use the same information sources and how they are integrated in aircraft systems. Abeloos et al (2000) suggest that surveillance data fusion of ADS-B and TCAS could improve airborne surveillance performance; and recommend that ASAS and TCAS data be presented on the same display. Zeitlin and Bonnemaison (2000) note the importance of TCAS/ASAS independence, and in particular stress that any loss of ASAS functions must not be detrimental to the functioning of TCAS, given its 'independent last resort' requirement. Hence, it is vital to understand the risks posed by 'common mode' failures and how they might be mitigated.

Estimates of 1 – Q(T) require simulation by aircrew, across the whole range of categories of encounter examined for P(T). Simulations of probability of detection for 'seeded errors' are necessary to test out the resilience of the system. The aim is to build confidence in the adaptability of the system rather than produce any kind of statistical proof. Resilience in this context is the number of safety barriers that are operative, but also has to provide assurance that system safety reaction times are sufficient. These seeded errors in simulation have to match the types of things that can happen, i.e. as generated by the error processes that lead to significant safety breaches. A further check would be to verify that all existing types of Airproxes are resolved.

How can it be known that the results of such simulations are 'right', or rather that they can deliver the kinds of statistical statements that are required? As noted in an earlier Section, the contributors to the special edition of 'Reliability Engineering and System Safety' in 1990 raised several wise concerns about simulation and the need for validation.

Some simple comments need making here. To start with, it is a question of confidence in results. Simulation can be an effective tool, but the trials must take place over comparatively long periods to eliminate unfamiliarity with new processes. Fortunately, ASAS as envisaged here requires aircraft/pilot simulation rather than workstation/controller simulations. Aircraft simulators can now be very realistic, as evidenced by the realistic behaviour of pilots in emergency scenarios. But the estimation of the BFP must also guard against aircrew being too aware that some kind of abnormal incident has been programmed. Thus, it is vital to test seeded 'blunders' in the context of a reasonably long stretch of normal operations rather than just present the aircrew with a high rate of blunders.

9. CONCLUSIONS

With Full Delegation Airborne Separation Assurance System (ASAS), separation control would be delegated to the (properly equipped) aircraft, i.e. aircraft pilots are responsible for their aircraft's separation from other flights. The aim is to try to identify a tangible work programme – rational and evidence based, and within the compass of known techniques – a framework that would prove safety.

Reasons for retaining the existing separation minima in an ASAS system have been put forward. The observed relative proportions of different types of Airprox suggest that, for the current system, comparatively large proportions of the Air

Traffic Services risk budget should be allocated to 'Reasonable Intent' risk (effectively 'right place on wrong flight path'). The key argument here is that mid-air collision in an ASAS environment will predominantly arise from this type of risk. Problems with the use of Probabilistic Risk Assessment with safety-critical events requiring probabilities to be estimated for 'human components' – Human Reliability Analysis – are reviewed.

The danger is the creation of 'over-elaborate' models – ones whose parameters cannot be reliably estimated from the data likely to be obtainable. Thus, the focus has to be on the simplest model that can be soundly based on available data. A simple 'Minimal Framework' Collision Risk Model is therefore constructed – which potentially can deliver practical results. It has three components:

> Significant Separation Breaches (SSB) – events when two aircraft have lost separation in all dimensions by a significant amount and action may be required to recover minimum separation;
>
> SSB/Collision Scaling Factor (SF) – the ratio of the long term average of collisions to SSBs if no safety defensive barriers were in place; and the
>
> Barrier Failure Probability (BFP) – the probability of failure of safety defensive barriers such as automatic warning systems.

The analysis here shows that these factors can be modelled by:

> SSBR – airspace geometry/traffic pattern plus human error simulation
>
> SF – kinematics of encounters
>
> BFP – equipment and human performance knowledge/simulations

Many of the building blocks for these already exist. This structure therefore enables these components to be integrated with specific additional work to be developed from well-specified experiments. But the challenge is to develop focused simulation tools and structured experiments to deliver quantitative outputs – results that are sufficiently convincing for the Rational Evidence Scrutiniser of Section 2. To reiterate the key point in that section's debate, this degree of 'confidence' can only come about through the creation of a compelling 'narrative' explanation, soundly based in theoretical understanding and empirical evidence, and open to challenge, checking and verification at every stage.

If this approach does not work, then what could?

ACKNOWLEDGEMENTS

This work was supported by a research grant by the Civil Aviation Authority's Safety Regulation Group (SRG). I would like to thank SRG staff and Ian Parker of National Air Traffic Services Ltd (NATS) for their comments on earlier drafts, and the referees for their cogent criticisms. I would also like to thank Ian Parker and Barry Kirwan of Eurocontrol for bringing me up to date on some current work on some strategic safety and human factors issues.

REFERENCES

Abeloos, A., Mulder, M., van Paasen, R. and Hoffman, E. (2000). Potential co-operations between the TCAS and the ASAS. International Conference on Human-Computer Interaction in Aeronautics (HCI-Aero 2000). Toulouse, 27-29 September 2000.http://www.eurocontrol.fr/projects/freer/archive/hci_aero.pdf.

Amalberti, R., (2001). The paradoxes of almost totally safe transportation systems, Safety Science 37, 109-126.

Barnett, A., (2000). Free-Flight and en route air safety: a first-order analysis, Operations Research 48(6), 833-845.

BASI, (1991). Limitations of the See-and-Avoid Principle. Australian Department of Transport and Communications.

Baumgartner, M. (2003). One safe sky for Europe – A revolution in European ATM. The Controller, July, 8-12.

Brooker, P. (2002a) Future Air Traffic Management – Passing the Key Tests, The Aeronautical Journal, 106(1058), 211-215.

Brooker, P. (2002b). Future Air Traffic Management: Quantitative En Route Safety Assessment Part 2 – New Approaches. Journal of the Institute of Navigation 55(3), 363-379.

Brooker, P. (2002c) Future Air Traffic Management Systems and Financial Decision-Making Constraints. Cranfield University Research Report PB/3/1/02, ISBN 1 861940 85 8 – to appear in 'Transportation' in a shortened form].

Brooker, P. (2003). Future Air Traffic Management: Strategy and Control Philosophy. The Aeronautical Journal, 107(October), 589-598.

Brooker, P. (2004a) Radar Inaccuracies and Mid-Air Collision Risk: Part 1 - a Dynamic Methodology – to appear in the 'Journal of the Institute of Navigation'.

Brooker, P. (2004b) Radar Inaccuracies and Mid-Air Collision Risk: Part 2 - En Route Radar Separation Minima – to appear in the 'Journal of the Institute of Navigation'.

Brooker, P. and Ingham, T. (1977). Target Levels of Safety for Controlled Airspace. CAA Paper 77002. CAA, London.

Brooker, P. and White, F. A. (1979). Minimum Navigation Performance Specification and Other Separation Variables in the North Atlantic Area. Journal of the Institute of Navigation, 32(3) 357-374.

The Chambers Dictionary, (1998). Chambers Harrap, Edinburgh.

Davies, E. H. and Sharpe, A. G. (1993). Review of the Target Level of Safety for NAT MNPS Airspace. CS Report 9301. NATS, London.

Dougherty, E. M. (1990). Human reliability analysis – where shouldst thou turn? Reliability Engineering and System Safety, 29, 283-299.

EC (European Commission) Transport RTD Programme, (1999). Emerald (Emerging RTD Activities of relevance to ATM Concept Definition) – Final Summary Report. http://www.cordis.lu/transport/src/emeraldrep.htm

Eurocontrol (1998). EATMS Validation Strategy Document. http://www.eurocontrol.int/eatmp/library/documents/EATMS_Validation_Strategy.pdf

Eurocontrol (2002a). ACAS brochure. http://www.nbaa.org/intl/acas2_training_brochure.pdf

Eurocontrol (2002b). Investigation of experience in modeling and determining separation minima. CARE-ASAS Activity 3. http://www.eurocontrol.int/care/asas/documentation/care-asas-a3-01-019.pdf

Eurocontrol SRC [Safety Regulation Commission] (2000a). Risk Assessment and Mitigation in ATM. Eurocontrol Safety Regulatory Requirement ESARR 4. Edition 1.0. Eurocontrol, Brussels.

Eurocontrol SRC (2000b). Safety Minima Study: Review of Existing Standards and Practices. SRC Doc 1, Eurocontrol, Brussels. http://www.eurocontrol.be/src/index.html

FAA/Eurocontrol (1998). A Concept Paper for Separation Safety Modeling An FAA/EUROCONTROL Cooperative Effort on Air Traffic Modeling for Separation Standards http://www.faa.gov/asd/ia-or/pdf/cpcomplete.pdf

FAA/Eurocontrol (2001). Cooperative R&D: Principles of Operation for the Use of Airborne Separation Assurance Systems (Version: 7.1) http://human-factors.arc.nasa.gov/ihi/documents/PO-ASAS.pdf

Foot, P.B. (1994). A Review of the Results of a Trial Hazard Analysis of Airspace Sectors 24 and 26S. CS Report 9427, Civil Aviation Authority, London.

Hacking, I. (1976). Logic of Statistical Inference, Cambridge University Press, Cambridge.

Hale, S. and Law, M. (1989). Simultaneous Operation of Conflict Alert and ACAS II in UK En-Route Airspace. DORA Report 8914, CAA.

Harrison, D. (1993). Results of ACAS II Safety Analysis. ICAO Secondary Surveillance Radar Improvements and Collision Avoidance Systems Panel (SICASP/5)

Harrison D. and Moek G. (1992). European Studies to Investigate the Feasibility of using 1000 ft Vertical Separation Minima above FL 290: Part II – Precision Data Analysis and Collision Risk Assessment. Journal of the Institute of Navigation, 45, 91-106.

Hoekstra, J. M., Van gent, R. N. H. W., Ruigrok, R. C., J. (2002). Designing for safety: the ‘free flight’ air traffic management concept. Reliability Engineering and System Safety 75, 215-232.

Hollnagel E., (1993). Human Reliability Analysis: Context and Control, Academic Press, London.

May, G. (1971). A Method for Predicting the Number of Near Mid-Air Collisions in a Defined Airspace. Journal of the Institute of Navigation, 24, 204-218.

Medawar, P., (1991). The Threat and the Glory, Oxford University Press.

Moek, G., ten Have, J. M., Harrison, D. and Cox, M. E. (1993). European Studies to Investigate the Feasibility of using 1000 ft feet Vertical Separation Minima above FL 290: Part III – Further Results and Overall Conclusions. Journal of the Institute of Navigation, 46, 245-261.

Moore, S. M., (1998). Comparison of Alerted and Visually Acquired Airborne Aircraft in a Complex Air Traffic Environment. Proceedings of the 1998 advances in aviation safety conference, SAE/P-321, 981205,held at Daytona Beach, Florida, April 6-8, 1998. SAE.

Moray, N., (1990). Dougherty's Dilemma and the One-sidedness of Human Reliability Analysis Reliability Engineering and System Safety, 29, 337-344.

NATO RTO Meeting Proceedings 32 (2001). The Human Factor in System Reliability – Is Human Performance Predictable? RTO-MP-032.

Quine, W. V., (1987). Quiddities: an intermittently philosophical Dictionary, Bellknap Press of Harvard University Press, Cambridge, Mass., USA.

Review of the General Concept of Separation Panel (RGCSP) (1995). Working Group A Meeting: Summary of Discussions and Conclusions. (1995). ICAO.

Sherali, H. D., Smith, J. C., Trani, A. A. and Sale, S. (2000). National Airspace\Occupancy and Conflict Analysis Models for Evaluating Scenarios under the Free-Flight Paradigm. Transportation Science 34(40). 321-336.

Shorrock, S. T. and Kirwan, B. (2002). Development and application of a human error identification tool for air traffic control. Applied Ergonomics 33 319–336.

Simon, H.A., (1982). Models of Bounded Rationality (Vols. 1 & 2). Cambridge, MA, MIT Press.

Simpson, R. W., (1998). Structuring Criteria for automated separation assurance, 2nd USA/Europe Air Traffic Management R&D Seminar. http://atm-seminar-98.eurocontrol.fr/finalpapers/track2/simpson.pdf

Swain, A. D., (1990). Human Reliability Analysis: Needs, Status, Trends and Limitations Reliability Engineering and System Safety, 29, 301-313.

THEATRE website, (2002). http://www.theatre.isdefe.es/home.sys.html?forum

Watson, S. R., (19945). The meaning of probability in probabilistic safety analysis Reliability Engineering and System Safety 45, 261-269.

Watson, S. R., (1995). Response to Yellman, T.W., and Murray, T. M., (1995). Comment on 'The meaning of probability in probabilistic safety analysis' Reliability Engineering and System Safety 49, 207-209.

Yellman, T.W., and Murray, T. M., (1995). Comment on ‘The meaning of probability in probabilistic safety analysis’ Reliability Engineering and System Safety 49, 201-205.

Zeitlin, A. D. and Bonnemaison, B., (2000). Managing Criticality of ASAS Applications. 3rd USA/Europe Air Traffic Management R&D Seminar. http://www.caasd.org/library/presentations/zeitlin.pdf

Zeitlin, A. D., (2001). Safety Assessments of ADS-B and ASAS, 4th USA/Europe Air Traffic Management R&D Seminar. http://www.mitre.org/support/papers/tech_papers_01/zeitlin_safetya/zeitlin_safety.pdf