Cranfield UNIVERSITY

AIR TRAFFIC MANAGEMENT ACCIDENT RISK
PART 1: THE LIMITS OF REALISTIC MODELLING

Peter Brooker

Version 1.0
6 May 2005
Cranfield Research Report PB/5/1/05

ISBN 1 861941 16 1

Contact details:

Professor Peter Brooker
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

CONTENTS

FIGURES

TABLES

AIR TRAFFIC MANAGEMENT ACCIDENT RISK
PART 1: THE LIMITS OF REALISTIC MODELLING

Peter Brooker

Cranfield University

"Mr Casaubon's…hope in immortality seemed to lean on the immortality of the still unwritten Key to all Mythologies"
'Middlemarch' by George Eliot

ABSTRACT

The prime goal of the Air Traffic Management (ATM) system is to control accident risk. Some key questions are posed, including: What do design safety targets really mean and imply for risk modelling? In what circumstances can future accident risk really be modelled with sufficient precision? If risk cannot be estimated with precision, then how is safety to be assured with traffic growth and operational/technical changes? This paper endeavours to answer these questions by an analysis of the nature of accidents, causal factors and practical collision risk modelling. The main theme is how best to combine sound safety evidence and real world hazard analysis in a coherent and systematic framework.

1. INTRODUCTION

The prime goal of the Air Traffic Management (ATM) system is to control accident risk. This leads to some important questions:

What are the essential ingredients for ATM systems to be designed safely?

How best can the lessons learned from accidents and incidents influence safe system design?

What do design safety targets really mean and imply for risk modelling?

In what circumstances can future accident risk really be modelled with sufficient precision?

If risk cannot be estimated with precision, then how is safety to be assured with traffic growth and operational/technical changes?

This paper endeavours to answer these questions in a coherent and systematic framework by an analysis of the nature of accidents, causal factors and practical collision risk modelling. The main theme is how best to combine sound safety evidence and real-world hazard analysis. It is the first of two papers. The second (Brooker, 2005d) applies this framework and toolkit to the important ESARR4 document and its supporting material, first through a critique and then by an attempted repair of ESARR4's methodology. [ESARR4 is the Eurocontrol Safety Regulatory Requirement Number 4 (Eurocontrol SRC, 2001): 'Risk Assessment and Mitigation in ATM'.]

One point needs to be made in advance. The analysis here will use some concepts from tort law. It is stressed that this is used because legal analysis has some admirable features when adapted to safety issues: legal analyses strive for clear definitions and logical frameworks. Those aspects of tort law concerned with blame and compensation are not relevant in any of the following, although some general comments are made about duty of care and safety responsibility. The word 'accident' used here just means an undesirable un-planned event, and words such as 'failure' have no connotations of blame.

This paper is standalone, but in some aspects it is a companion document to earlier papers, in particular Brooker (2004b), which examines the use of quantitative risk targets, and Brooker (2005b), which focuses on learning lessons from hazardous incidents.

2. WHAT IS THE AIR TRAFFIC MANAGEMENT SAFETY SYSTEM?

The question 'What is the 'Air Traffic Management Safety System' that provides Safety?' actually covers several related questions. These include:

What is Air Traffic Management (ATM)?

How is ATM different from Air Traffic Control (ATC)?

What is meant by 'System' in 'ATM System'?

Given an ATM System, where are the responsibilities for safety?

How is the system designed for safety and its performance improved?

There are probably no absolutely correct answers to these kinds of questions, but some kinds of answers may be more useful or more prudent than others. The focus here is on the safety of commercial transport flights carrying passengers and/or freight receiving an ATC service, focusing mainly, but not exclusively, on mid-air collision risk. The most valuable understanding of the ATM System's characteristics is taken to be one that best facilitates safety improvements.

'Safety' itself has many meanings. Here it will be taken to be something measured by risk rates to people – aircraft accidents involving deaths (Brooker, 2004b). What kinds of accidents are definitely not ATM System safety issues? Three categories are security, aircraft flight and pilot flying:

> Security covers deaths from hijacking, sabotage and other acts of terrorism, where there is an intent to harm. Obviously, the air travel system needs security defences, but these are additional to and distinct from safety defences, so risks from security failures should be assessed separately from aviation safety.
>
> Accidents arising purely from aircraft flight are those in which the aircraft systems fail in some catastrophic fashion. Examples would be loss of engine power, major structural failure of the aircraft control surfaces or engine mountings, shutdown of cockpit electrical systems. In these circumstances, the aircrew could probably not prevent an aircraft crash, nor would any intervention by ATC be likely to alleviate the situation.
>
> A 'pilot flying' accident would be one arising from aircrew action at a critical stage. A pilot might mis-remember the critical aircraft speeds on takeoff and hence fail to get airborne. A pilot might have a heart attack during an instrument approach and dive the aircraft into the ground. Once again, no kind of intervention by ATC could be guaranteed to alleviate the situation.

Aircraft flight and pilot flying are therefore aviation safety risks but not ATM System risks. Thus, the categorisation of safety is as in Figure 1.

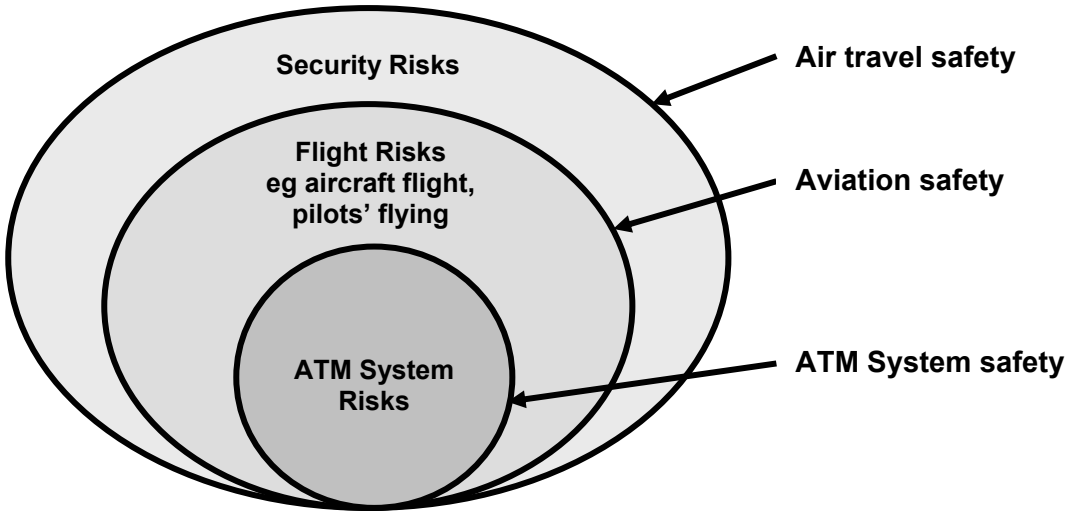

Figure 1. Differences between air travel risk categories

What is inside the ATM System safety circle above? First, all ATC functions should be in the circle. ATC has information about the likely paths of other flights, and therefore can pass instructions to the pilot that will ensure safe passage on taxiways, runways and in flight. For present purposes, ATC covers all the activities from the start of a flight to its completion that involve the determination of the aircraft's flightpath, including planning activities and flow management (even though some of these tasks could take place before the aircraft is in motion). Some aspects of this may be related to scheduling operations rather than safety, but they are still generally held to be an ATC function. There may some boundary definition issues, eg with airline operational staff and airport ground staff. A test question might be: "Is the movement of the aircraft being planned, monitored or changed with the movement of other aircraft in mind?"

As far as the aircrew are concerned, what happens 'inside ATC' is a black box. The aircrew get instructions and, so long as they appear sensible and prudent, the aircrew do not care if they derive from judgements by human beings or the outputs of computer processing. The aircrew know that ATC has a duty of care and is subject to safety regulation, so that the internal workings of ATC will make sense in safety terms, and hence controllers' instructions should be obeyed.

ATC is carried out by what are often termed ANSPs (Air Navigation Service Providers), such as National Air Traffic Services (NATS) in the UK. ATC uses several kinds of technical entity: communications, navigation, surveillance (CNS), and data processing DP. These are surely all part of the ATM System? The kind of test questions that should be posed are: 'Would these entities be there if the ATM System were not, and would the ATM System require them to be in place in order to deliver the levels of safety required?'

Data processing (DP) links the CNS functions. Communication systems, both air-ground and between controllers and ATC centres, including both voice and electronic data transmission, are an integral part of ATC: effective ATC requires the controller to communicate with the pilot. Surveillance is also now an integral part of ATC: radars are there so that controllers can see position information on screens. Without it, how could controllers do their job?

Navigation is different. Controllers do not navigate aircraft, but ANSPs do provide some ground aids for aircraft, while others are independent on-board systems (eg altimetry for vertical navigation). But ground aids can also be provided by military organisations that are separate from civil ATC provision, while independent public/private bodies could provide satellite-based systems. Being part of the ATM System does not necessarily imply anything about ownership. So is navigation part of the 'ATM System'? Yes, it is, because it would be very difficult to run a safe system in any meaningful sense of the phrase if aircraft flightpaths did not match their ATC instructions. If aircraft could not navigate well, mid-air collisions might be much more frequent. Note that a collision attributable to a gross navigational equipment failure could only occur if that equipment had been deemed by providers and regulators to be 'fit for purpose', ie they had accepted it into the ATM System.

What about airborne collision avoidance systems on the aircraft (ACAS)? They are not provided as part of ATC. They exist to compensate for an ATC or pilot manoeuvring failure of some kind. Are they part of the ATM System? This raises large questions about the nature of the ATM System safety – see Brooker (2005a). Note that ACAS is not intrinsic to flying the aircraft; it is there to protect against conflicts with other air traffic. ACAS has been installed systematically, by agreement between States in ICAO, and has to meet appropriate performance requirements. They are not there through a decision by ANSPs alone, but because it was believed by States that the system needed further management to ensure that it delivered the necessary safety. This surely adds up to ACAS being part of the ATM System.

ACAS is (obviously) not part of ground-based ATC. ANSPs do not supply, provide or maintain ACAS equipment. It is States, advised by their aviation safety regulators, which agreed to introduce ACAS, working within the larger context of Eurocontrol, the European Union and ICAO. In the UK, the 'State' would be the Department for Transport (DfT) and the regulator is the Civil Aviation Authority (CAA) Safety Regulation Group (SRG) (other countries have different kinds of structures, but the responsibilities and professional tasks are broadly similar). Both the State and its regulator therefore play a part in the ATM System.

A test question would be to ask if accidents could happen more frequently if DfT and SRG do not do their jobs properly. The answer is yes. For example, suppose DfT failed to support the introduction of safety improvement equipment; or if SRG were to fail to set out the right kinds of rules and regulations for its performance or fails to monitor its operational usage. The regulators cannot act passively: it is not enough to promulgate regulations in the office and read an ANSP's safety case documents; thought must be given to the safety consequences of regulations and appropriate inspections of real operational practices.

DfT and SRG are not of course ANSPs. But they have responsibilities for ensuring that the UK ATM System is up-to-date and delivers the necessary safety. Similarly, in the UK, the Directorate of Airspace Policy has safety responsibilities. Thus, if airspace is poorly configured, so that (eg) controller workload is badly affected, or if regulations are unclear, so that accidents happen because of confused decision-making, then there could be increased risks. The point is that these bodies have some measure of control over the design of ATM system and its operations, and this control carries with it a duty of care (compare tort legislation on fault, eg, Williams and Hepple (1976), page 92 et seq). All the participants in ATM system design and operation have a duty of care to make reasonable, skillful and proper use of relevant evidence, and generally to apply intelligence and foresight. Thus, the participants must demonstrate that they meet the standards for aviation safety professionals.

Taking all the above into account, the ATM Safety System has the components in Figure 2, which just shows national dimensions.

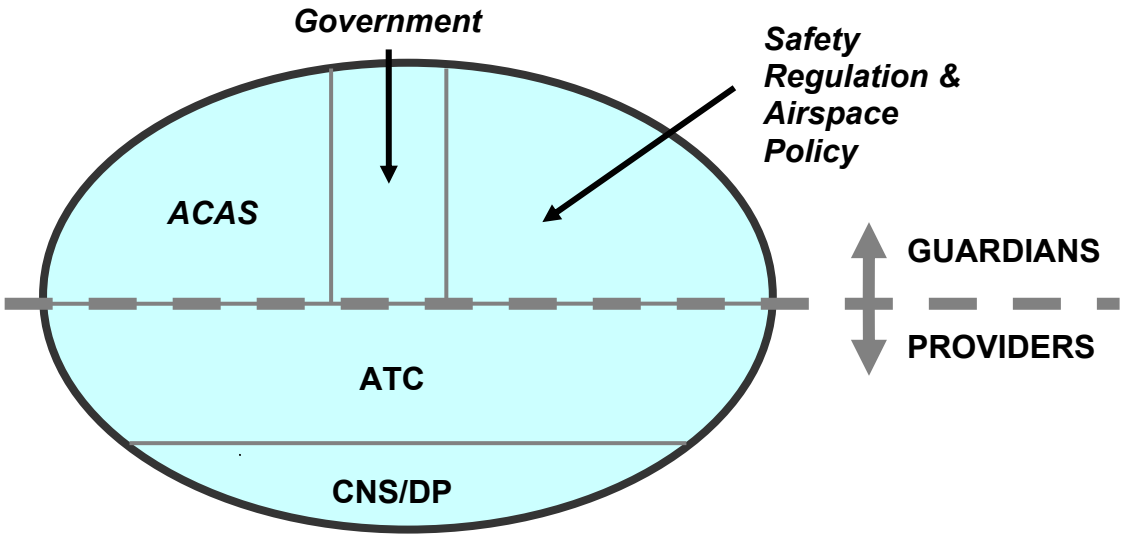

Figure 2. Components of ATM System safety

Hence, the ATM System is much more than the provision of a ground based ATC service. The thick dashed line in the diagram divides its elements into 'Guardians' and 'Providers'. The Guardians inter alia regulate the Providers – but who or what regulates the Guardians? The answer would appear to be that the Guardians make their deliberations openly and expose their arguments to technical criticism.

Do all the parts of Figure 2 have to be in the ATM System? The crucial point is that their interactions and dependencies are vital to the creation of something that delivers safety – which is why it is better to consider them together as an integrated package rather than try to break them apart. One test question is to ask if a particular component's safety responsibilities could – in principle – be passed to another element in Figure 2. If it cannot be passed, then that component is surely an integral element of ATM System safety?

3. ACCIDENT RISKS

What are the essential ingredients for ATM systems to be designed safely? How best can lessons learned from accidents and incidents influence safe system design?

The measurement of past safety performance is important, but it is not necessarily the best tool for improving performance. Information about achievements does not always tell us about how to make things better in future. Real safety improvement relies on an increasing understanding of causal factors for the most probable future accidents *plus* the right tools for dealing with these factors. ATM is highly successful in safety terms, so the sadly ironic problem is that accidents are rare and individual events. This is because a sequence of additional safety defences is now in place to limit accident risk. The mechanisms that occur in ATM accidents tend to bring together combinations of circumstances that are unlikely to reoccur, even over many decades.

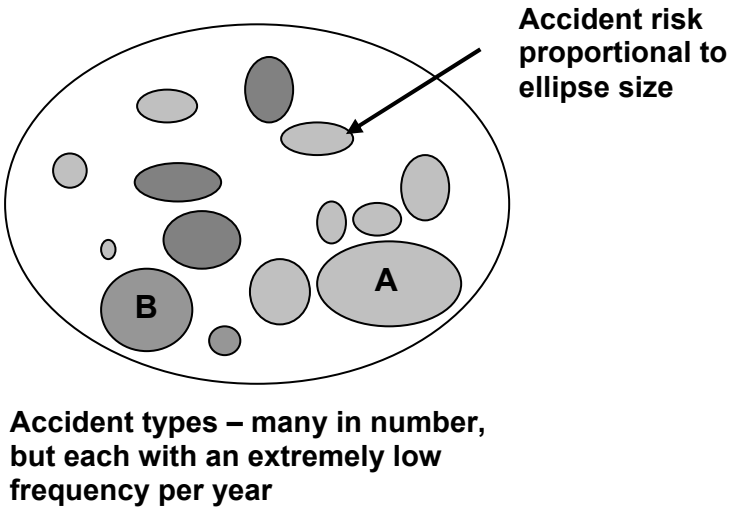

Figure 3. Accident Types and Risks at a particular time

This nature of ATM accident Types and risks is illustrated in Figure 3. The areas of ellipses represent the probabilities of accidents of a particular 'Type'. Nearby accident ellipses represent accident Types that have some kind of closeness; that are 'in the same family'. Each of these accident Types occur extremely rarely (ie in reality the corresponding ellipses have tiny dimensions), thanks to the successes of ATM's safety defences, continued high quality training and operational discipline, and indeed the lessons learned from past accidents. The most probable accident shown in Figure 3 is of Type A, but it may not happen for many years – and perhaps a Type B accident could occur much earlier, just by chance.

Suppose a Type A accident does occur. If this is in any way a new variety of accident in system terms, the whole industry will work hard to find ways of preventing a Type A accident from reoccurring. Its lessons will be added to the safety knowledge base.

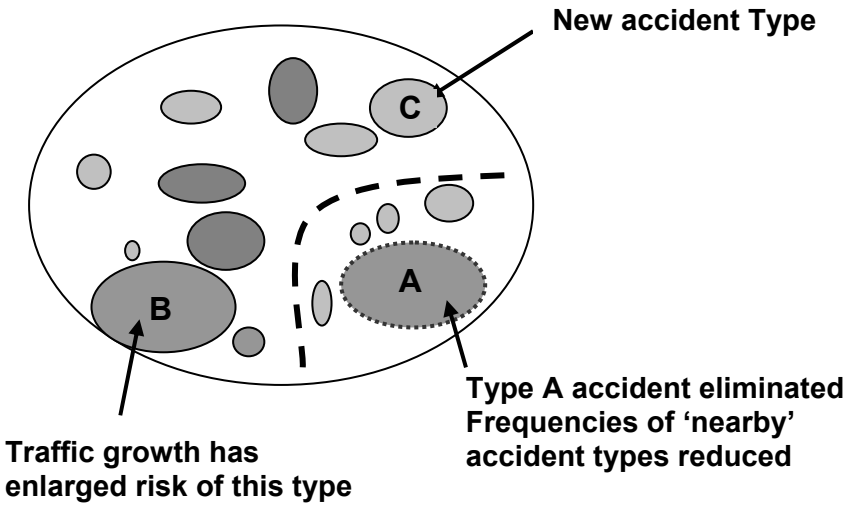

Figure 4. Accident Types and Risks at some future time after a Type A accident

At some future time, if safety improvement actions corresponding to Type A are completely successful, Type A accidents will be eliminated from the set of accident Types, so the area of its ellipse will shrink to zero (Figure 4). Moreover, the solutions to the Type A accident will tend to have beneficial effects of nearby accident Types, so the frequencies of these Types will also be reduced. Hence, in the Figure these nearby ellipses will shrink, to indicate a reduced probability. But it may be that Type A accidents cannot be eliminated (which leads on to the topic of safety targets examined in a later section). This may because there are no straightforward technical ways of dealing with some weak spots in safety defences. Or there may not be any 'reasonably practicable' ways of eliminating them, except possibly over several years.

But traffic growth between the two time periods may have increased the probability of some other kinds of accident Type – eg Type B. There may also be new types of accident – Type C in the Figure. The latter could arise because of some novel equipment or operational concept. Again, the frequency of any new accident Types would be extremely small, given the intense professional scrutiny of such innovations. But some level of risk might get through the analysis process – how is this to be constrained?

Compared with the first time period, the accident rate in the second time period would be reduced by the feedback actions and increased by the effects of any new accident Types. Given the maturity of the ATM system and continuing attention to safety, this would be expected to deliver a progressively improving risk level, the 'natural' rate of safety improvement. It should be noted that this safety evolution is an international process, not a national or regional one. Safety analysts, particularly in the more developed nations with manufacturing and system development capabilities, use all the worldwide information they can obtain to improve their ATM system.

Thus, safety analysts have to rely on general knowledge about the nature of failures, the effectiveness of safety defences and the information contained in incidents about what actually happens. Experienced analysts will no doubt have in their minds an estimate of what these ellipses and their sizes might be. These estimates are based on a combination of general understanding of system failure modes; lessons from incidents modelling of potential failure modes, extrapolation, evidence of incidents – and also the analysts' gifts of imagination.

4. CAUSATION AND NESS TESTS

The previous section has rather glibly discussed accident Types without explaining what such a thing might be or how to categorise such abstractions. An understanding of the nature of these Types is essential if the process and limitations of risk modelling are to be understood. The discussion here uses a concept known as a NESS test, where the abbreviation is for 'Necessary Element of a Sufficient Set' (Wright, 1988 and 2001).

At the heart of any understanding and modelling of accidents has to be a methodology for expressing their causation. Only if one knows what causes

accidents and how frequently these causes are in place can any prediction be made of the frequency of future accidents. Causation is an inherently complex topic, but the first question asked about an accident is what caused it (the second being 'how do we prevent a similar accident in future?').

The literature on causation covers psychology, engineering, logic/philosophy and legal aspects. Most of the ATM literature focuses on the first two of these; the shift here is to try to use some of the legal thinking on causes, which is itself backed up by philosophical approaches. To repeat the warning in the Introduction, the interest here is on finding ways of describing causation in accidents not the tort law concerned with responsibility, blame and compensation. The focus here is the search for the meaning of "causally relevant condition" and to find tests by which a condition can be judged causally relevant.

Key references on causation start with Hume (1777). The next major advances in thinking about complex causes were made by Mill (1904). In particular, Mill observed that there may be many different distinct sets of conditions that are each sufficient to bring about the effect, which means that there is no unique sufficient set, the idea of *plurality of causes*. Thus, many (different) causal combinations could potentially produce mid-air collisions. The NESS test approach used in the following originated with Mackie (1980), Hart and Honoré (1985), and Honoré (2001). The NESS test was developed by Wright (1988 and 2001), with critical input from Fumerton and Kress (2001). The discussion here concentrates on the use of the NESS test rather than its underpinning or complexities, and makes particular use of Honoré (2001).

First, two key concepts need to be briefly defined: necessary and sufficient. Again, these are intrinsically complex concepts (eg Honoré, 1995). For two sequential events or states A and B, the definitions would be on the lines of:

> *A* is *necessary* for *B*, if the non-occurrence of *A* guarantees the non-occurrence of *B*.

Thus, the second event would not have occurred had it not been for the former. [NB: A is the 'antecedent' and B the 'consequent'.] .

> *A* is *sufficient* for *B* if the occurrence of *A* guarantees the occurrence of *B*.

Thus a *sufficient condition* is one which brings about the event, but where such an effect could also have followed from other factors.

Accident investigations generally (but sometimes implicitly) use the 'but for' causation test. In determining whether a particular system failure caused the accident, it is asked whether the accident would have occurred *but for* that failure. Thus, the actual state of affairs is being compared with a hypothetical state of affairs – what would have happened had that particular failure not occurred. The 'but for' test is not a conclusive test of causation. It is often described as a *negative* test, serving the purpose of screening out factors that were in fact irrelevant to the outcome.

Mackie's (1980) work on causation added understanding about sets of causal conditions. Each condition in such a set should be identified as a necessary and non-redundant part of the set, and where all the conditions are at hand the consequence follows with necessity. There may be several different sufficient sets, so for any particular set of conditions is not *actually* necessary for the consequence – there may be an alternative set (or sets) of conditions that will produce the same effect.

The definition Wright proposes, the Necessary Element of a Sufficient Set (NESS) test, stipulates that:

> *A particular condition was a cause of a specific result if and only if it was a necessary element of a set of antecedent actual conditions that was sufficient for the occurrence of the result*.

Wright's test is based on the idea that a fully described causal law would list all the conditions that are sufficient for a certain effect. Irrelevant conditions are eliminated with the requirement that only those antecedent conditions that are necessary for the sufficiency of the set are included. The NESS test combines this with Mill's idea of a plurality of causes, so that there is no unique sufficient set.

Under the NESS test, the hypothetical question is slightly different, but similar, to the 'but for' test. Would the remaining set of conditions, if the condition in question had not been at hand, still have produced the consequence? Thus, the NESS test asks whether the condition being tested was a necessary element of the set of conditions sufficient to bring about the result. The simplest way of performing the test is to, hypothetically, eliminate the condition in question from the set and consider, against the known applicable causal generalisations, whether the effect would still occur. But note that the possibility of simply eliminating a condition is not always available – in other cases it is necessary to replace the missing condition with another one, based on knowledge about what the world 'should be'.

A simple non-aviation illustration of the NESS test would be person A negligently starting a fire, and then later, as the fire is going out naturally, person B accidentally pouring petrol on it. The fire spreads and causes significant damage to person C's property. A's action in negligently starting the fire is a NESS cause of the damage to affecting C, as is the action of B in pouring petrol on the fire as it was going out. It is not possible to create a subset of actual conditions sufficient for the result without including the actions of *both* A and B. So A and B are both necessarily part of the set of sufficient causes for the harm to C.

Thus, to summarise, the definition of the NESS test says that to cause an accident means to complete a set of conditions sufficient to bring the accident about. This clearer concept of causation helps eliminate some problems resulting from a diffuse concept of causation. The NESS test is not a panacea, but it facilitates identification of the correct factors rather than unessential or irrelevant elements.

The NESS test needs to be supplemented by clear thinking about the degree of remoteness and probability that can be allowed for when something is being assessed as a causal factor. For example, suppose a lack of airline check-in staff

results in a delay to many of the passengers on a flight, in an otherwise normal traffic situation, which means that that the aircraft misses its flow management slot, takes off later, and is then involved in a mid-air collision. This staff problem should not be treated as a NESS cause of the accident or the passengers' deaths. The point is that, on the basis of information available at the time, the probability of an aircraft accident was not substantially increased by the delay to the takeoff arising from the passenger handling problems.

Thus, the decision on remoteness can be converted into a question about whether or not the system designers/operators could reasonably have foreseen the harm that could result from this failure to operate in accord with the standard operating requirements in force at that time. This is actually a potentially very difficult question for system designers, because their prime task is to envisage the consequence of both standard and non-standard aviation operations coupled with various kinds of system failure. What they can 'reasonably foresee' is a much tougher test compared with the average member of the public. In particular, it is the task of system designers to change standard requirements.

The delayed passengers example is actually a straightforward case, because it is easy to see that the airline managers did not have the kind of information available to them that would have indicated a higher probability. Indeed, the accident's causes would very probably be events occurring after the aircraft had taken off. As this scenario has assumed generally normal traffic, it is difficult to see how a failure of ATM system design could be labelled as a cause, unless the system designers could somehow have constructed some kind of predictive model demonstrating markedly higher risk. But, again, such a model would have to predict reliably the failures arising from pilot/controller decisions post-takeoff.

So what is the connection between accident Types and NESS tests? The answer is that an accident Type can be defined here as a distinct set of conditions that are sufficient to bring about the accident. The accident Type is thus a minimal set of causes as constructed by applications of the NESS test – the 'NESS causes'.

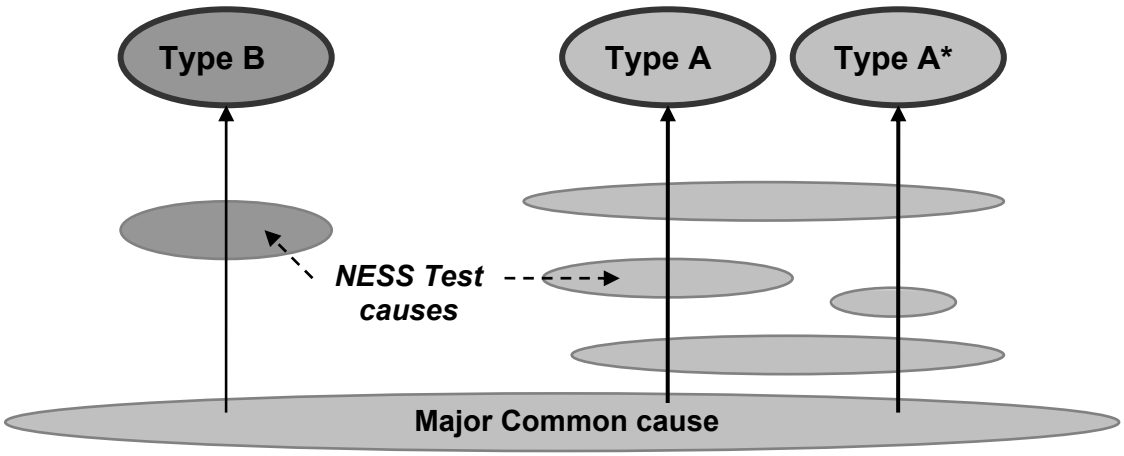

Figure 5. Different Accident Types and NESS causes

The relation between accident Types and Ness tests is illustrated in Figure 5. There are three accident Types shown: A and A*, which are 'nearby', and B, which is of a very different kind. The NESS causes are shown as shaded ellipses leading up to the Type risk ellipses. Again, the size of the NESS ellipses represents some kind of risk statement; it will be taken later here to be a conditional probably of that kind of event or action occurring given the previous events or actions.

A and A* are nearby because they share all but one of their NESS causes. B is distant from them because it has just one shared cause, the bottom one, which is labelled as a major Common cause. This common cause might be an inherent flaw in airspace design. B in the Figure is, all other things being equal, a simpler accident Type than either A or A*, just because there it has fewer NESS causes.

5. ACCIDENTS AND INCIDENTS

Over the history of aviation, the likelihood of a mid-air collision being of Figure 5's Type A, with several causes, has tended to increase, while the chance of it being of Type B, with few causes has markedly decreased. This is simply because system defences have been added that eliminate or substantially reduce the probability of the 'one or two causes' accidents.

Thus, the first airliner mid-air collision was in 1922, between a French Farman Goliath and British de Havilland DH18 at Thieuloy-Saint-Antoine in France. The aircraft were flying in opposite directions using exactly the same ground route, in bad visibility (Flight, 1922). Thus, unsafe route design was a major Common cause.
In 1956, there was a mid-air collision over the Grand Canyon in the USA between a Constellation and a DC-7. The causal factors identified were (ICAO, 1956):

> "The pilots did not see each other in time to avoid the collision. It is not possible to determine why the pilots did not see each other, but the evidence suggests that it resulted from any one or a combination of the following factors: 1) Intervening clouds reducing time for visual separation; 2) Visual limitations due to cockpit visibility, and; 3) Preoccupation with normal cockpit duties; 4) Preoccupation with matters unrelated to cockpit duties such as attempting to provide the passengers with a more scenic view of the Grand Canyon area; 5) Physiological limits to human vision reducing the time opportunity to see and avoid the other aircraft, or; 6) Insufficiency of en-route air traffic advisory information due to inadequacy of facilities and lack of personnel in air traffic control."

This accident led to a complete overhaul of the USA aviation system, in particular the expansion of controlled airspace and faster introduction of secondary radars and flight data processing; and indeed can be seen as the genesis of airborne collision avoidance systems. For present purposes, the key feature is the increase in the number of causal factors and the wider system design and management issues.

The most recent airliner mid-air collision in European airspace was the Überlingen accident (BFU, 2004). Nunes and Laursen (2004) tried to establish a causal chain for the accident. Their list of 'Contributing Factors' in the Überlingen accident includes:

(a) Single Man Operations

(b) Downgraded Radar [STCA]

(c) Dual Frequency Responsibility

(d) Phone System

(e) ACAS

(f) Corporate Culture

Assessing these as NESS causes is an interesting exercise. Each has to be examined very carefully. For example, the ACAS cause is not actually a failure of the ACAS equipment but that the nature of its alerts was not communicated automatically to the controller. Moreover, the Corporate Culture heading covers the fact that 'the integration of ACAS (TCAS II) into the system aviation was insufficient and did not correspond in all points with the system philosophy' (BFU, 2004).

In the present context, the main message drawn by Nunes and Laursen is that the Überlingen tragedy was a consequence of multiple system failures – operational, managerial, technical and regulatory. This accident had several NESS causes.

Accidents are rare, but incidents are less so. What should be judged to be a 'serious' incident, given the formal description of accident Types attempted above? The word incident can be used in several different ways. For ATM mid-air collisions it is often taken to be an event in which the separation minima required between aircraft is breached (eg see the National Air Traffic Services SSE scheme sketched in Neil et al (2003) and the FAA/Eurocontrol (1998) report on separation minima). But is this the most useful viewpoint of 'serious' in safety terms? On the NESS test approach, an accident is an occurrence in which there is a sufficient set of causes, so a serious incident might best be judged as one in which just one or possibly two of these causes were absent. Thus, an incident is seen as being an 'incomplete accident', where the addition of just one or two causal factors would have produced the accident comprised of those NESS cause elements.

A simple non-aviation analogy of incident seriousness would be a motorcyclist who falls off because of a bad stretch of road surface. The chance of the motorcyclist being killed by a following vehicle is dependent on the traffic at that time. It might happen that there is no other traffic for some minutes, so the rider can remount safety. But if the condition of the surface could be equally poor for the next motorcyclist, when other traffic might be much greater, then this represents a serious incident in terms of the need for road improvement.

Thus, on the incomplete accident view of an incident, the most serious ones would be those in which just one extra causal factor would complete the accident, and for which the conditional probability for that factor was the largest, ie they would be the most probable consequence of the occurrence of these completing NESS causes. [Brooker (2005b) discusses this issue starting from different premises.] These are the most 'useful' serious incidents, in that they provide the best information about potential future accidents, ie the probable sizes of the ellipses in Figure 3. This also implies that reducing the likelihood of Common causes is a key ingredient to

improving safety, because they can be completed to become accidents in so many different ways.

Some important examples of incomplete accidents can be found in Airprox reports (UKAB, 1999-). Table 1 is an extract from Brooker (2005c). The first column is the Airprox number (the first four digits are the year); the Summary is an edited version of the Airprox report text; and the right hand columns are the horizontal (H) and vertical (V) distances at the closest point of approach.

Airprox number	Summary	H Nm	V feet
1999127	Controller had issued a descent clearance that would have led aircraft 1 to descend through the level of aircraft 2, which he had inadvertently not taken into account.	0	1100
1999200	Controller did not take aircraft 1 into account when he descended aircraft 2.	4.5	400
2000032	Controller gave 'erroneously and essentially unforced descent instruction' to aircraft 2.	2	700
2001069	Controller allowed aircraft 1 to climb to the level that he had cleared aircraft 2 to fly at, without coordination.	2.8	700

Table 1. Examples of Commercial Air Transport Airproxes with ATC descending/climbing an aircraft into another's path: 'Outside [STCA] parameters'

The key point about the incidents in Table 1 is that STCA did not alert the controller: the manoeuvre took place to quickly for it to operate – the meaning of the 'Outside [STCA] parameters' in the figure title. ACAS did operate. ACAS, plus the density and geometry of traffic in the airspace at the time, were the operational safety defences. Thus, these actually correspond to a small number of NESS causes rather than a many-cause accident such as Überlingen.

In the safety literature, these incidents would generally be considered as 'Errors of Commission', although some would actually fall into the class of 'Erroneous Execution'. This is in contrast with a failure of a controller to detect/act on a large deviation from the planned flightpath – which would usually be classed as an 'Error of Omission' (eg see Hollnagel, 2000). Relevant definitions are:

- Error of omission: lack of action, so system/component/function status quo is preserved rather than (necessarily) changed;
- Error of commission: one that changes the system to an unsafe state.

Errors of commission when aircraft are in proximity are obviously potentially very hazardous. Thus, it is vital to find ways of preventing them, of detecting them, and of ensuring that the remaining safety defences would be effective.

6. SAFETY TARGETS, ACCIDENT TYPES AND COLLISION RISK

The previous paragraphs start to enable the possibility of answering the questions: What do design safety targets really mean and imply for risk modelling? In what circumstances can future accident risk really be modelled with sufficient precision?

How is completeness of accident Types to be ensured – how is it possible ensure that all abnormal modes are included? If these desires cannot be met, then how is safety to be assured with traffic growth and operational and technical changes?

This section first outlines some background on important features of Target Levels of Safety (TLS): this is a summary of Brooker (2004b). A TLS is a design hurdle, a quantified risk level that a system should – ie be designed to – deliver. It has to provide assurance that the future system – with changes – is as safe, preferably safer, than the present one. The TLS relates to '*total system design*…the implication is that all types of failure, mechanical, procedural and human, which generate a risk of collision will be accounted for' (Brooker and Ingham, 1977). A TLS covers *all* aviation-related causes. However, it does not usually attempt to cover the consequences of terrorism or criminal behaviour, ie where there is an intent or willingness to cause an accident (although the literature has not always been clear on this).

A TLS appropriate for accidents arising from mid-air collisions has been developed since the 1970s. It is usually derived by taking historical accident rates, which show a progressive reduction over time, and extrapolating forward, thus getting tighter and tighter over time. The TLS is measured in fatal aircraft accidents, ie accidents in which at least one person in the aircraft was killed, per so many aircraft flying hours. The recent ICAO figure of 1.5×10^{-8} fatal aircraft accidents per flying hour (RGCSP, 1995) is the rate corresponding to mid-air collisions – for any reason and in any spatial dimension – in en route flight in controlled airspace.

Mid-air collisions are now rare. Because of this, it is not possible to estimate the current accident rate in Europe with great statistical confidence. To validate with statistical confidence a rate of 1.5×10^{-8} fatal aircraft accidents per flying hour would require of the order of 10^9 observed hours. But the annual European flying hours are only of the order of 2×10^7. Estimates of the future accident rate (an Actual Level of Safety – ALS), given traffic growth and new operational and technical features, therefore rely on risk modelling. Risk modelling has to rely on an understanding of the causes represented in accident Types, which necessarily includes extrapolation of present system features, in particular human performance and failure rates. The ALS has to cover all the ellipse blobs in Figures 3 and 4.

Quantitative ATM risk modelling has been carried out using a variety of mathematical, computational and simulation models. The range of models is now very large. An extremely useful overview of the field is set out in FAA/Eurocontrol (1998), which inter alia provides 40 pages of annotated bibliography. This report's title includes the phrase 'separation safety modeling', and much of the ATM collision risk work has been concerned with setting separation minima between aircraft: how far they should be kept apart procedurally (eg flight levels) and by ATC actions (eg distance between aircraft being vectored). These minima fulfil many functions (eg see Brooker, 2004a), but they are essentially the key ATM 'system safety control' parameters.

An analogy for separation minima is a road speed limit. The road accident rate could be reduced to virtually zero if all vehicles went very slowly. But society recognises

that people want to travel quickly, so is prepared to trade off the statistical expectation of some accidents for a 'fair' speed. Speed is obviously not the only factor – some drivers are drunk and others are just reckless. Nor can the speed limit be enforced at every point on every road. These factors add up to the present speed limit and 10 to 20 people being killed every day in the UK. ATM separation minima correspond to a far tighter safety regime in terms of the restrictions on pilots, the existence of ATC, conflict detection systems and safety monitoring. There is still a trade off, but it is made explicit by safety management and risk modelling – and it produces uneconomic flight paths for some aircraft.

What are the causes of – future – mid-air collisions? Again, the subject is intrinsically very complex. FAA/Eurocontrol (1998) provides about 25 pages outlining (!) factors that can potentially affect separation safety; Brooker (2005c) examines UK data on Airproxes for commercial air transport aircraft using UK controlled airspace with a radar service, in which STCA and ACAS are functioning properly. [NB: an ACAS alert is in some sense a failure of controller-provided separation.] A simple general picture is shown schematically in Figure 6.

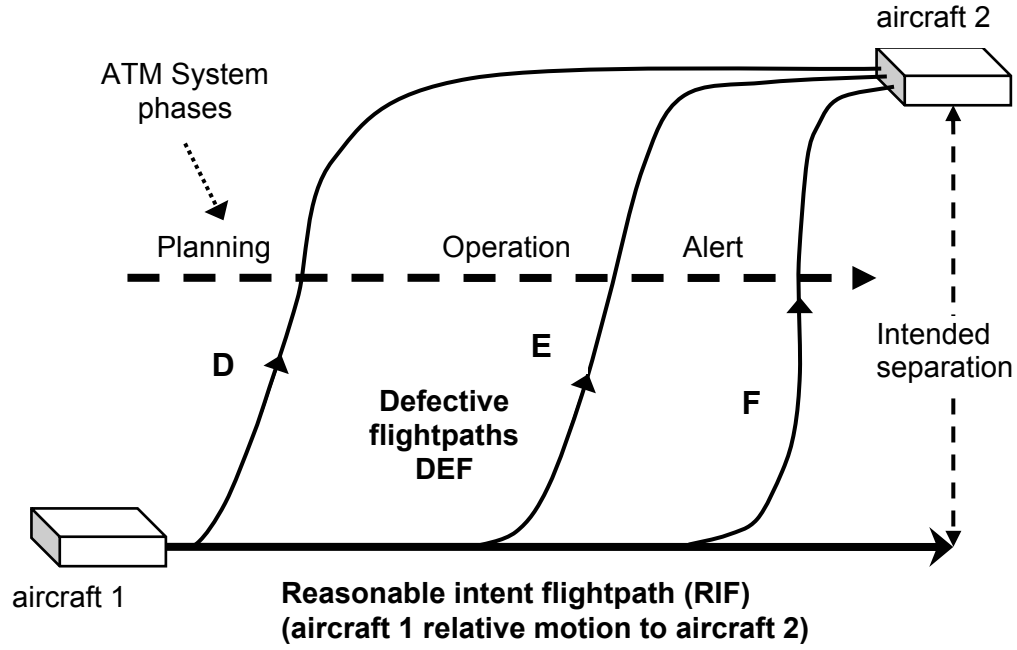

Figure 6. Schematic of defective flightpaths leading to mid-air collision

The scenario shown in Figure 6 is aircraft 1 suffering a mid-air collision with aircraft 2, both shown as boxes. The picture is in the frame of reference in which aircraft 2 is at rest, ie it shows the relative flightpath of aircraft 1. The orientation of the aircraft is left open – at the outset, they could be horizontally or vertically separated. Aircraft 1 follows the 'reasonable intent' flightpath (RIF), shown here as a straight line but in reality usually much more complex. By reasonable intent is meant the following (Brooker 2002):

> Reasonable Intent An inference usually made 'after the event'. The reasonable intent flightpath here is what a competent controller would have considered a reasonable (albeit perhaps not perfect) course of action; the

pilot's decisions and actions were what other pilots would have judged suitable practice (albeit perhaps not the ideal decisions).

It must be stressed that RIFs are not always static concepts. Some flightpaths have to be tactical in nature, because a complete set of failsafe routeings cannot be provided into the immediate future of the flights. Environmental conditions, eg the wind vector and way that the aircraft is being flown, can mean that what was predicted to be a safe flightpath has to be changed tactically by the controller in order to ensure continued safety.

A defective flightpath (DEF) is a gross deviation from the RIF. This implies – a crucial point that will be discussed further – that to monitor/measure DEFs the analyst needs to know the corresponding RIF). A DEF would include people's misjudgements and blunders as normally understood. It would also cover gross loss of 'Position Integrity': operationally extreme ('out of tolerance') errors on radar, GPS, altimetry; measurements lost, markedly inadequate display information; signals being corrupted or lost. The huge improvements over recent decades have eliminated many of these technical problems.

Associated with the RIF is an intended (or implied) separation between the two aircraft at their planned closest point. A collision occurs if this separation is 'eaten up' by some kind of DEF. This might be when two aircraft pass each other or when two aircraft supposed to be flying in-trail ('station keeping') lose the required separation. Figure 6 correlates the evolution of the reasonable intent flightpath with the ATM system phases of the aircraft's flight. The three phases noted here, Planning, Operation and Alert are and explained further later (for more detail, see Brooker (2005). The words used have normal meanings; thus, Planning refers to the entry of the aircraft into the airspace, Operations covers the work of controllers (eg monitoring the aircraft's flight), and Alert covers STCA and ACAS, including the actions of controllers and pilots in response to such alerts.

Three examples of defective flightpaths are shown: marked **D**, **E** and **F**:

- **D** DEF occurs very early on, in Planning phase, eg with aircraft being put on the wrong routeing; then controller monitoring and alerts all fail to get aircraft back onto RIF
- **E** DEF occurs during Operational phase, eg controller neglects to ensure separation, and the alerts fail to get aircraft back to RIF
- **F** DEF occurs when aircraft are near to their closest point of approach, so that some or all of the alerts are not effective (compare the examples in Table 1).

The Überlingen tragedy probably falls into the **D** category. Most UK observed Airproxes involving commercial air transport aircraft are incomplete versions of **E** and **F** (eg see the analysis in Brooker, 2005c). Note also that the DEFs shown in the Figure correspond to the last defective flightpath for a flight. Previous defective segments of flightpath that have been successfully returned to the RIF complicate the picture (compare Brooker, 2005b). These previous defective flightpaths and corrections would not usually be NESS causes of an accident, because their existence would not have been necessary for it to occur.

Thus, the first ingredient for a mid-air collision is some initiating event to produce a DEF – call it iDEF. Not all such events will produce a collision. The next ingredient is a combination of failures of the system's remaining safety defences, so the aircraft is not finally returned to the RIF. Again, this may be one NESS cause or several depending how far to the right the DEF occurs in Figure 6 – the 'Right Hand' (RH) defences. The third ingredient is that the aircrafts' closest approach to each other is sufficiently near for the possibility of a collision, which depends on how large the aircraft are and their relative velocities. [Note that a close approach might also be converted into a mid-air collision if the RH defences do not play their role effectively – an 'induced collision'.] The final factor also has to include elements to describe the density, dimensions and velocities of aircraft. This combination of 'traffic factors' and aircraft dimension/velocity – a 'kinematic scaler' – is really a scaling factor for the airspace sub-system under consideration rather than a 'cause' in the normal sense of the word. For example, it has to cover the fact that aircraft at the same nominal altitude are, through 'chance' altimetry errors, separated vertically.

The traffic factors reduction in risk is more than 'providence'. Large, cautious distances are fed into the system design, so that only gross deviations are safety-significant and so that ATC can detect and act in time. Note that, if the aircraft are not flying in 'formal' structure routeings, eg parallel tracks or crossings on fixed route systems, then the traffic factors + kinematic scaler have to be some kind of probabilistic average of potential risk configurations. The nature of the DEF will affect relative velocity factors.

With each NESS cause can be associated a conditional probability. This is the chance that, given what has gone before, this particular cause will be operative. Thus, in a situation in which a controller has climbed an aircraft into the path of another flight, the next NESS cause would be some kind of failure of the STCA/controller combination to provide a safe resolution to the encounter. This probability value is not a fixed value for all encounters – it will depend on the relative positions of the aircraft and their closing speeds AND the context of the controller's tasks at that time. How can such probabilities be estimated with precision?

The mid-air collision rate CR (which can be stated as a number per so many system flying hours) can therefore be written as expression *A*:

$$\mathbf{CR = \Sigma\ R(iDEF) \times P(RH\ failures) \times (Traffic\ factors) \times (Kinematic\ scaler)} \qquad A$$

Here:

R	rate (eg) per hour
P	probability
iDEF	particular initial DEF event
Σ	summation over all possible iDEFs
RH failures	all the safety defences to right of iDEF fail
Traffic factors	appropriate densities of conflicting aircraft iDEF
Kinematic scaler	appropriate aircraft size and velocities combination for iDEF

The analytical equation *A* for CR above is an exact one. It does no more than spell out the mechanisms by which collisions logically have to occur. The hard problem is how to populate the parameters in expression *A* with sensible numbers.

What would be sensible numbers? In the present context they must be numbers for which there is reasonable statistical evidence, ie which are known to a degree of precision. They should be capable of validation in the largest sense, ie that an independent, competent person could realistically carry out the same measurements and/or would produce very similar extrapolations on measured data. What can reasonably be modelled with precision – what is knowable? A probability that cannot be independently verified has little real existence.

This parameter population process must achieve high standards. The following extract, from studies on the Space Shuttle (NASA, 2005), is just as pertinent to civil aviation safety:

> "Validation – Determination that an item meets its intended purpose in its operating environment. For models/analysis tools, design environments and simulations, validation is the determination that the item accurately reflects the subject being modelled…Central to the safe and reliable conduct of high-risk complex technical endeavours is rigorous and consistent understanding of, and adherence to, [validation and other] terms and the processes they describe.
>
> This understanding and adherence also applies to methods leading to the end state (ie models and analysis tools utilized during validation). As an example, if one is to assert 'validation has been accomplished through probabilistic analysis', the analysis must rest upon fundamental mathematical principles and undergo unflinching rigor."

These are very demanding criteria. They place limits on the kinds of calculation that can be deemed to meet the standard required. They raise issues about the use of a technique such as TOPAZ [TOPAZ is the acronym for 'Traffic Organization & Perturbation AnalyZer]. Descriptions of TOPAZ are given in Blom et al (2003) and ARIBA (1999). TOPAZ uses a range of sophisticated mathematical techniques. It does not appear to correspond to simple physical pictures of the consequences of erroneous events/failures. Because it is to some extent a commercial product, it is not possible to track how the model works through following simple real-life examples. It is not possible to verify that the model as it might be used in practice ensures any kind of completeness in hazard analysis terms. It is not possible to verify that the navigational performance and human factors-related probability distributions used match the range of evidence that can be gathered. The published material does not appear to present detailed information on data, distribution extrapolations or sampling fluctuations.

TOPAZ techniques may well be of use, particularly in understanding how different human factors aspects might contribute to risk. But they do not appear to provide decision-makers with a robust, explicit and verifiable way of answering a key safety question.

7. RISK ESTIMATION

So how are the rates R and probabilities P to be estimated? A starting point is to break down the collision risk calculation into different types of ATC-planned separation. To understand how collisions might occur, it is necessary to spell out how ATC is structured and acts to prevent collisions in these different circumstances. Proximate aircraft will be kept separated in different ways: some will be vectored horizontally by ATC; others will cross each other's flightpath with vertical separation; other pairs will be kept longitudinally separated (eg prior to being sequenced for landing); and some aircraft will be kept safely separated by being put on parallel tracks. Moreover, there are different types of ATC process in different phases of flight, eg compare terminal area and en route ATC regarding the tailoring of the flightpath for appropriate sequencing. Call these different types of ATC process 'proximity categories'.

In each proximity category, ATC should always be ensuring a particular kind of separation in at least one specific x, y or z dimension – but note that it cannot be assumed that all aircraft will enter the sub-airspace of interest 'separation compliant'. The accident Types corresponding to these categories will therefore tend to have some NESS causes in common. Each proximity category will have corresponding traffic factors and kinematic scalers. The total collision risk CR will be a summation of the collision risks in the proximity categories weighted by flying hours/movements, ie expression *B*:

CR = Σ Σ R(iDEFc) x P(RH failures) x (Traffic factors) x (Kinematic scaler) *B*

where iDEFc now represents an initial DEF event in the proximity category c.

The key to estimating collision risk is an understanding of the causal nature underlying the rates R and probabilities P in the general CR expression. A simple picture of the Figure 6 process would be a four-stage process shown in Figure 7.

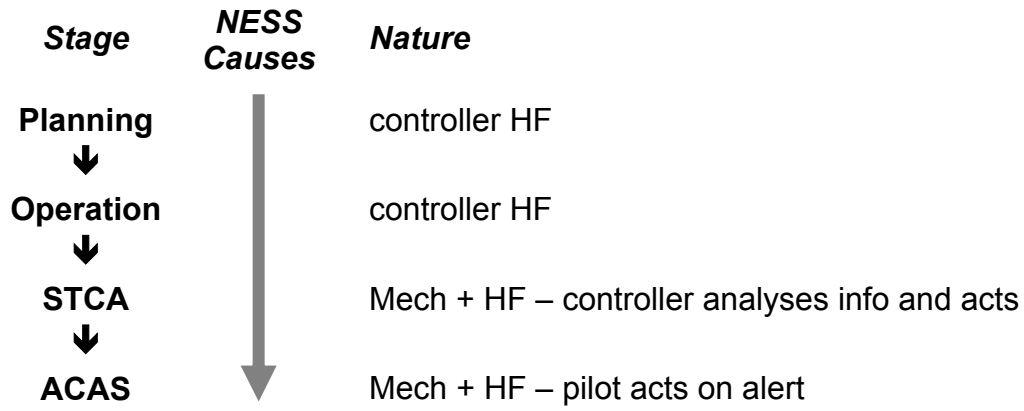

Figure 7. Simple four stage ATM processes for errors and recoveries

In Figure 7, HF stands for human factors and Mech for mechanistic (eg through simple computer extrapolation of aircraft paths based on radar data). Mech + HF means Mech combined with HF; thus, a controller gets an STCA, analyses the circumstances and then makes a judgement about the correct action. The Mech components in the first two stages would generally be expected to be second order

compared with the HF aspects. The NESS Causes for a particular accident Type will be spread across the stages according to the accident Type being considered.

The HF rates and probabilities will depend on the ATC context leading to the iDEF and the subsequent recovery stages. They are intrinsically dependent on 'soft' HF issues, such as selection, training, acceptable workload and safety culture. Thee will be regulatory influences, eg consider the Überlingen tragedy causes referenced above, where the probability of a successful intervention via TCAS was 'modified'.

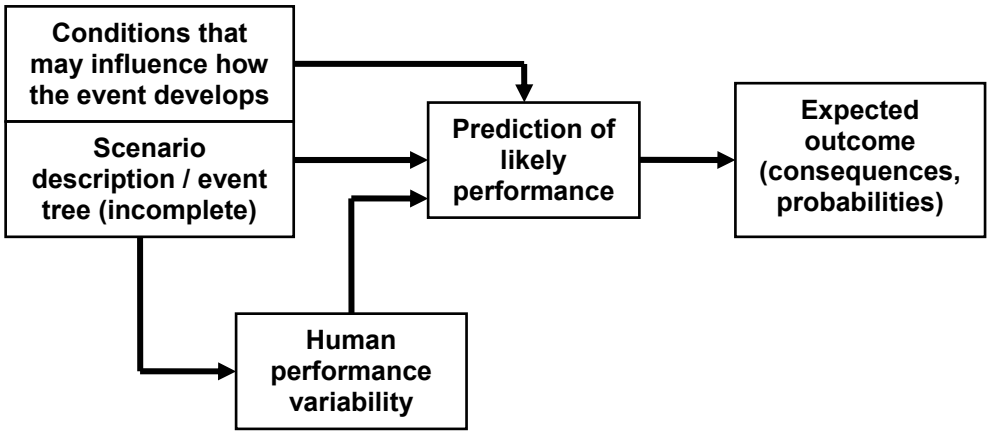

Figure 8. Risk dependencies in performance prediction (adapted from Hollnagel, 2000)

The problem of ATC context is illustrated in Figure 8. To quote Hollnagel (2000):

> "The overall purpose of performance prediction is to describe how a scenario may possibly develop, given the existing working conditions. In many cases the representation of a scenario only provides the basic structure of the events but leaves out the detailed conditions that may influence how an event develops. In order to make the prediction, the scenario description must therefore be supplemented by information about the conditions or factors that can influence the propagation of events. One of these is the variability of human performance, which in itself depends on the general performance conditions—including the previous developments … Performance prediction must therefore describe the likely context *before* it goes on to consider the actions that may occur".

The inherent predictability problems are well summarised by Dougherty (1997):

- the context of performance is variable;
- performance itself is variable;
- performance and its context are too complicated, with too many parameters.

The last is a lack of knowledge, which may be reduced over time, but only if the right kind of data is collected.

Consider the first two stages in Figure 7. Suppose an initial DEF occurs and is recovered. What would be the rate of ATC generating such a DEF and the probability that it is recovered by ATC? [NB: it is important here to distinguish

between non-DEF deviations, the normal variations in track, ie the equivalent of flight technical error, and a DEF, which requires special recovery attention.] But where is the data for this to come from? – it is unlikely that this is routinely recorded. Is there good data on the frequency with which controllers operate tactically or use particular separation minima? Is it really possible to detect systematically DEF planning defects that are being corrected in operation? Are DEFs monitored and recorded by the ANSP when there is no apparent risk of mid-air collision?

The rate of DEFs has to be linked to the causal factors associated with the rate at which trained and experienced pilots and controllers generate DEFs or certain kinds of error leading to a DEF. If the risk assessment were being carried out for a system in which there are changes to HF processes, it would be necessary to understand the causal structure for the elements (Figure 8) affected by such changes. Exactly the same kinds of questions need to be posed for the operational recovery stage.

The data situation is sometimes better for the Alert stage, simply because alert events are more likely to be recorded systematically. The UK has its Airprox reporting system, which provides information about STCA and ACAS events. Thus, there is the probability of real knowledge of DEFs detected by STCA and ACAS. Something can therefore be said about probabilities of success of the Alert phase by examining STCA/TCAS II RA events, eg Brooker (2005).

7.1 Probabilistic Safety Assessment

Risk modelling in most safety critical industries is accomplished through Probabilistic Safety Assessment – PSA (other terms, such as Probabilistic Risk Analysis, PRA and Quantitative Risk Assessment, QRA, are also used). The use of PSA is often suggested as appropriate for ATC collision risk estimation – but most often by people who have not attempted to use it for this purpose in real life. What are the issues and problems?

The key point of the PSA approach is that the risk of accidents is estimated by analysing the whole sequences of events that could produce an accident: the 'causal chain' (eg Apostolakis, 2004). 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 potentially be determined by observations of the performance of that particular sub-system (although simulations and 'expert judgement' are also used). Thus, a good PSA estimate requires a sufficiently detailed understanding of failure modes and engineering characteristics.

But complex engineering systems, particularly ATM, generally contain people as an intrinsic part of the 'safety-delivery' operation. These people have to assess information and act in certain ways when confronted by specific system states, eg a controller seeing two aircraft in unplanned proximity must issue instructions to restore safe separation. Thus, these kinds of events necessarily require probabilities to be estimated for 'human components' – the task of Human Reliability Analysis (HRA).

It is inherently difficult to produce estimates of event frequency for infrequent occurrences. The modelled PSA chain may well be appropriate, but the difficulty is in 'populating' it with relevant data. An important example is Foot (1994), a trial PSA of UK en route air traffic operations. This study demonstrates the combinatorial explosion in fault tree complexity – there are a very large number of potentially hazardous system states – and hence the requirement to estimate an equally large number of failure mode parameters.

The traditional way of getting around the problem of the inherent uncertainty in PSA 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 ACAS – of which more later) and overestimating failure rates, then the problem is resolved. 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.

Probably the most compelling example of these calculation problems is set out DNV (1997), which dealt with estimating the safe spacing of P-RNAV parallel routes. It is a lengthy, detailed and thoughtful piece of work using best practice PSA techniques. Considerable efforts were made by DNV to try to ensure completeness of description and quantification of significant critical hazards; and major efforts were made to gather and assess relevant human error and performance data. DNV (1997) produced collision risk estimates for parallel routes taking account of ATC intervention. Unfortunately, the uncertainties associated with the predicted risk covered what DNV referred to as a large "grey" band (eg for same direction traffic, the safe spacing was estimated as between 5Nm and 15Nm).

The DNV work illustrates the problems of PSA very clearly. ATC safety analysis necessarily requires sequences and probabilities to be estimated for failure events involving 'human components'. It is very difficult to produce estimates of these generally infrequent events, particularly for errors of commission, and a collection of 'cautious' assumptions produces over-pessimistic – not practically usable – risk estimates. To quote the subsequent DNV (2003) report:

> "Using hazard analysis and associated techniques to estimate the tails has been tried in the past with inconclusive results. The uncertainties associated with such approaches are large and the benefits for this particular scenario relative to extrapolating known data are unclear."

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 particularly good analyses of the issues. More recent references about the problems of HRA are Embrey (2004) (in particularly referring to work by Kirwan (1994) on HRA accuracy) and Apostolakis (2004)).

Dougherty (1990) 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

Moray (1990) commented:

> "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."

These methodological concerns have not been resolved in the last 15 years; eg to quote Apostolakis (2004)

> "Several items that are either not handled well or not at all by current QRAs are:
>
> *Human errors during accident conditions*. For an accident in progress, we can distinguish between errors of omission (the crew fails to take prescribed actions) and errors of commission (the crew does something that worsens the situation). These errors, especially those of commission, are not handled well and research efforts are underway to improve the situation…It is also important to point out that experience has shown that the crews often become innovative during accidents and use unusual means for mitigation. These human actions are not modeled in QRAs.
>
> *Safety culture*. When asked, managers of hazardous activities or facilities say that they put safety first. Unfortunately, experience shows that this is not always the case. While it is relatively easy to ascribe an accident that has occurred to a bad safety culture, the fact is that defining indicators of a good or bad safety culture in a predictive way remains elusive. QRAs certainly do not include the influence of culture on crew behavior and one can make a good argument that they will not do so for a very long time, if ever."

Perhaps the best overview evidence of the limited usefulness of PSA/HRA methods in ATM is given by FAA/Eurocontrol (1998). This was a joint FAA and Eurocontrol Organization effort to 'model the impact on safety resulting from changing required separation minima and introducing new technologies in controlled, domestic airspace'. Some 30 recognised safety experts developed this authoritative report. It consists of 178 pages, including an extensive bibliography of all the techniques successfully used in ATM risk assessment worldwide. Less than 1½ pages of the

document were concerned with PSA/HRA; without any mention of major successes in tackling real-world ATM questions.

A collision risk PSA incorporating a HRA is thus a complex calculation process that is likely to produce usable answers only at some indefinite point in the future. Can a simpler model framework be constructed that is soundly based on available data?

7.2 Collision Risk Modelling (CRM)

As evidenced in FAA/Eurocontrol (1998), the great bulk of practically useful accident risk estimations use the simplest model that can be soundly based, ie whose parameters can be reliably estimated from the data likely to be obtainable. The phrase Collision Risk Model (CRM) will be used here to describe analytical frameworks on which are 'hung' empirical/statistical data about DEF rates and failure probabilities. Reich (1966) was the first generally accepted model of this type; a review of such models is set out in FAA/Eurocontrol (1998); some successful applications are listed later here.

A CRM is essentially developed by compressing the stages in Figures 6 and 7. It must not lose 'risk content' from the basic CR equation. To lose such content would be a failure to ensure completeness of risk estimation.

The first step is to restrict the analysis to a particular well-specified proximity category c (eg a loss of vertical separation for aircraft on a fixed route structure). But this must not be done too finely; otherwise it will become very difficult to assess what should be the corresponding safety target. The ideal would be to add all these categories together – but the following will show why this is intrinsically difficult. This is mainly done to restrict the complexity of the failure and probabilities, traffic factors, and Kinematic scaler. So the expression *B* becomes:

$$\mathbf{CRc = \Sigma\ R(iDEFc) \times P(RH\ failures) \times (Traffic\ factors) \times (Kinematic\ scaler)} \qquad C$$

The next step is to move away from the initial DEF, iDEFc, to the final DEF, fDEFc. This is the rate of DEFs before the final recovery response to urgent action by the controller, which may have involved automatic alerts. If a DEF were returned to the appropriate RIF by 'normal' controller operational action, then it would not be counted as a fDEFc. In terms of the CR expression, the first two terms are changed to:

$$\mathbf{CRc = \Sigma\ R(fDEFc) \times P(RHf\ failures) \times (Traffic\ factors) \times (Kinematic\ scaler)} \qquad D$$

where the second term again refers to the probability of failure to resolve, but this time after the final DEF. Again, note that this recasting of the equation has not lost any risk content. The final step is to re-label and reorder the equation, putting at the right hand side the probabilities relating to controller and pilot actions

$$\mathbf{CRc = \Sigma\ fDEFc\ Performance\ rate \times Traffic\ factors \times Kinematic\ scaler \times Pilot\ \&\ controller\ action} \qquad E$$

This final expression *E* is of the form used in the Reich and similar models. Reich (1966) and FAA/Eurocontrol (1998) show examples of the formal algebraic expressions, in particular the kinds of expressions to be expected for the Traffic

factors and Kinematic scaler components. The only issue to be resolved is what is included the final action probability class: interventions by controllers have been included in some models (eg Brooker and Lloyd, 1978); STCA has been considered as a possibility in others (eg DNV, 2003) and argued for (Brooker, 2004c); policy about incorporating ACAS has been raised by Brooker (2005a).

The practical use of a model of the expression *E* form has to rest on a number of assumptions and caveats:

(i) All the elements in this equation must potentially be measured or estimated by appropriate extrapolations of statistical distributions.

(ii) The HF elements in the early generation of DEFs have been hidden in the expression. They are within a 'black box'. The final DEF rate asks about the rate at which defective outputs are generated. Hence, the rare event frequency is obtained by reasoned factoring of infrequent but measurable events, rather than through full understanding of mechanisms and component probabilities. [This is the main deficiency of CRM in terms of risk estimation for markedly changed operational concepts.]

(iii) To know if there has been a DEF it is necessary to know what the corresponding RIF is. This may be a deterministic issue, eg deviating from a parallel route system, or a stochastic one, eg the relative orientations of aircraft being vectored horizontally.

(iv) To know the rate of <u>all</u> DEF occurrences per hour of operation, it is necessary to have some kind of measurement process. This could be routine, (eg through radar deviation monitoring), or through special measurement programmes, or by realistic simulations, or through rigorous reporting by controllers/pilots of gross deviations from the RIF. Data is crucial: it must be available or gettable. It is usually necessary to extrapolate DEF characteristics from this source data.

(v) A category c may contain several components with different conflict geometries, eg dependent on whether the DEF was caused by a data entry error or a miscommunication of information. These different geometries may correspond to different relative velocities at closest approach, ie the kinematic factors could be markedly different.

The above are actually quite restrictive constraints, which limit the possible application of this kind of (potentially) precise CRM to specific kinds of ATM sub-system. The practical successes of a CRM approach (mainly en route but some for airport issues; most, but not all, based on variants of the Reich model) include:

- VOR-defined routes (ICAO, 1976)
- Longitudinal NAT separation (Brooker and Lloyd, 1978)
- NAT Track System (Brooker and White, 1979)
- Radar separation (Sharpe, 1991)
- Precision Runway Monitor (FAA, 1991)
- RVSM (Harrison and Moek, 1992)

- P-RNAV parallel routes (DNV, 1997 and 2003)

To stress again, these are real examples of practical successes that have been used in key decision-making.

7.3 Loosely- versus Tightly-coupled Models

A useful way of thinking about potential ATM accidents is to construct two broad categories according to the kind of technological and human system structures that are being employed to ensure safety. These are called tightly- and loosely-coupled models. These terms originated with Weick (1976), and were subsequently used by Perrow (1984) for the purpose of analysing accidents.

Perrow in fact defined two important dimensions: interactive complexity and loose/tight coupling:

> Interactive complexity refers to the presence of unfamiliar or unplanned and unexpected sequences of events in a system, either not visible or not immediately comprehensible.
>
> A tightly-coupled system is highly interdependent, with each part of the system being tightly linked to many other parts, so a change in one part can rapidly affect the status of other parts. So tightly-coupled systems respond quickly to perturbations – but this response may be disastrous.
>
> Loosely-coupled systems have less tight or fewer links between their parts, so they are able to absorb failures or unplanned behaviour without destabilization.

A loosely-coupled system allows some 'play' in the system stabilizing (negative) feedback loops—a little over correction, followed by some under correction. Loose systems are more adaptable, have more tolerance for error, but can have much longer reaction times. If what happens in one part has little impact on another part, or if everything happens slowly, eg on the scale of human thinking times, the system is not tightly-coupled. Loosely-coupled systems tend to be open and continually interacting with the outside environment.

A tightly-coupled design generally uses traditional engineering methods, with bits of electronic kit, aircraft construction, software, etc. Tightly-coupled systems can survive failures, but only if that kind of failure has been anticipated and provided for in the original design. Designers of tightly-coupled systems must therefore invest effort and thought into anticipating failure modes and providing safety features to permit survival and recovery. In contrast, loosely-coupled systems tend accommodate failures through adaptive responses.

Perrow's particular concern is with safety-critical systems that have both interactive complexity and tight coupling. In such systems, an apparently trivial incident can potentially cascade in unpredictable ways that cannot be remedied, and hence produce severe consequences. However, as is made very clear in Marais et al (2004), ATM does not in fact fall into this Perrow category. In general, much of ATM system design is deliberately de-coupled in order to increase safety. In particular,

large minimum separations are planned between aircraft, so that mistakes by controllers can be remedied; hence loosely-coupled. This is in a system containing independent and engineering-redundant safety defensive layers (Brooker, 2002).

Some sub-systems of the ATM system are designed to be tightly-coupled. The operation of these kinds of tightly-coupled designs can usually be modelled quantitatively. These models can be validated against what happens in the real world. Thus, the sub-system acts in a 'programmable' or routine fashion (with specific designated functions). The key element is that the range of expressed 'failure modes' is comparatively limited and well defined.

These kinds of approximately tightly-coupled systems would include navigation of well-defined route systems, vertical separation and ILS. In such cases, Human Factors 'failures' need to be sufficiently regular in nature to permit a simple accident model to be used, ie of expression *E* form. For example, it may be possible to measure the frequency of a straightforward error in inputting the correct data into an aircraft computer and put this in the model as a DEF rate.

In contrast, loosely-coupled ATM sub-systems would include the pilot flying the aircraft away from airport runways and ATC/pilot interactions in sectors. Loosely-coupled ATM designs use much more complex information sources. For example, the controller's job requires visualization and situational awareness skills.

Figure 9, taken from Brooker (2005), shows, in a very abstract and simplified fashion, the transit of a typical flight in ATM system terms. The three pre-operational – Planning – layers have been grouped together because they are highly related, eg separation minima depend on suitable equipment being available, while the controller has to work within the safety constraints using the equipment. The Operation Layers cover the activities of pilots and controllers while the flight is in progress. The Alert Layer is the ground and air protection enabled by STCA and ACAS, on which the controller/pilot will act. (The figure ignores the reality of complex feedback loops.)

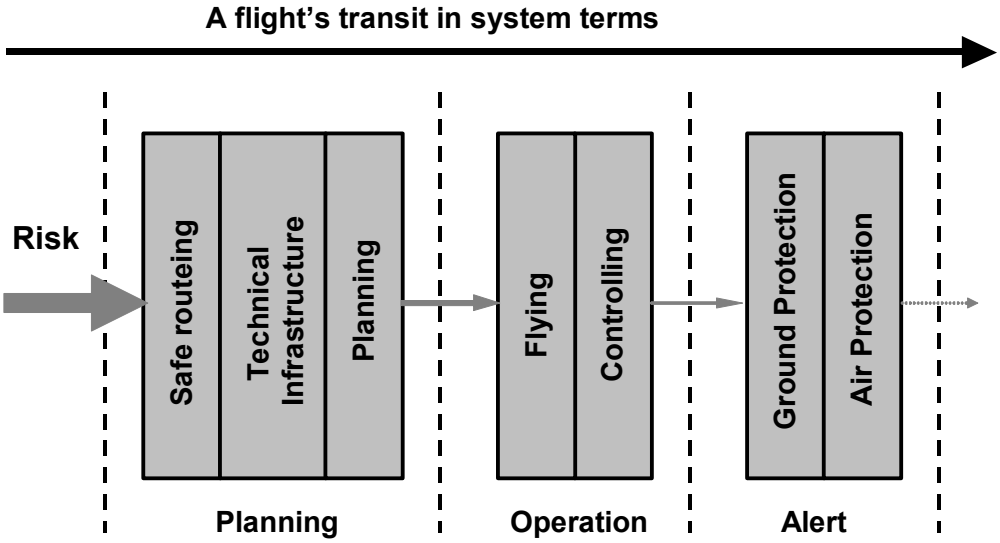

Figure 9. The ATM system layers – highly simplified, without 'loops'

But what do these layers actually accomplish in terms of system risks? The answer is that they act systematically to reduce mid-air collision risk. The purpose of the system layers is to reduce the 'end product risk'. A formal structure is imposed by the Planning Layers, next the Operation Layer should eliminate inherent conflicts (but note that the Planning Layer does not produce conflict free paths from departure to landing), and then the Alert Layer warns the controller and pilot about impending conflicts. Thus, the risk arrow in the diagram should shrink at every stage; in other words, each defensive layer scales down the probability of a potentially hazardous situation.

This is the only kind of quantitative risk model that is practically feasible for a loosely-coupled sub-system (or Accident Type category). An attempt at PSA will fail for the same reasons already noted: too many options, too large a potential for adaptive response and flexibility, and too many probabilities to estimate. Loosely-coupled sub-systems have 'slack' in time or space, and the potential kinds of HF interaction are many in number. This also means that an attempt at a 'precise' CRM will also fail – too few of the CRM conditions will be met.

It is possible to carry out risk calculations for some loosely-coupled accident Types. Examples are Brooker (2002, 2003, 2004a). The basic idea is to use selected Airprox data as indicating a DEF rate and then scale down by density and aircraft flightpath geometry. There are obvious issues: what is an appropriate set of Airproxes? Could under-reporting be a major problem? Is enough known about the geometries and relative velocities of the aircraft at closest approach?

These factors all tend to degrade the precision of the estimate, making it both rough and probably pessimistic if a sequence of cautious modelling assumptions has had to be made. One way of dealing with this is to include specific factors for the beneficial effects of STCA and ACAS. The policy implications of this are discussed in Brooker (2005a) (in particular the section on Risk Assessment Over-pessimism). The key point is that, if the defensive barrier benefits of STCA/ACAS are excluded from ATM system safety calculations, then this puts an extra burden on risk

estimation, in that the calculations will tend to be unjustifiably over-pessimistic about the value of new concepts.

Safety for loosely-coupled operational systems is improved – purportedly "to meet safety targets" – by an on-going process of safety feedback plus the introduction of additional safety-related defensive layers and engineering redundancies, eg GPWS, STCA, ACAS, error-free FMCS databases, etc. The key thing to ensure safety is that the rest of the ATM safety layers described above work effectively enough to produce the necessary corrective action. For example, there is a need to focus attention on circumstances and geometries when STCA and ACAS do not provide large amounts of extra protection or when the geometries/velocities mean that they induce risk.

To summarise the key features of called tightly- and loosely-coupled models in the present context:

> Tightly-coupled models – accident risk is a function of specific failures, eg gross navigational errors or a restricted set of Human Factor failures occurring comparatively regularly. Risk can be numerically quantified in terms of a limited number of key failure modes using CRM.
>
> Loosely-coupled models – safety is provided through a structure of defensive layers: risks occur if these layers perform poorly and do not filter out potentially hazardous situations. Risk can only be roughly numerically quantified, based on past defensive layer performance.

8. CONCLUSIONS

Some important questions are posed here about ATM safety, including: What do design safety targets really mean and imply for risk modelling? In what circumstances can future accident risk really be modelled with sufficient precision? If risk cannot be estimated with precision, then how is safety to be assured with traffic growth and operational/technical changes?

These questions have been addressed by an analysis of the nature of accidents, causal factors and practical collision risk modelling. The underlying theme is how best to combine sound safety evidence and real-world hazard analysis in a coherent and systematic framework.

The essential ingredients for answering the questions – a mental toolkit – include:

- Clarity about the nature of a safe ATM System, which is much more than the provision of a ground based ATC service, eg includes regulatory factors.
- Rationale for accident Types and probability of occurrence.
- Concept of 'NESS causes' describing causation of different accident Types.
- Tendency for the number of NESS causes in accidents to increase over time, because of improvements and added safety defences.

- General theory of collision risk modelling, based on gross flightpath deviations (DEF) and recovery from ‘reasonable intent’ flightpaths (RIF), noting that these can occur at different stages in ATM system phases. .
- Analysis of the inherent difficulty in mapping causes with large human factor components onto measurable gross deviations.
- Evidence that PSA/HRA will generally be unsuccessful in estimating ATM risks.
- Criteria for collision risk modelling to provide estimates with usable precision.
- Distinction between loosely- versus tightly-coupled ATM system models, with examples of approximate collision risk models based on Airprox data.

An important conclusion is that there are intrinsic limits to circumstances where realistic quantitative modelling is feasible. For increased coverage by such models, a number of criteria would have to be met, eg in terms of the data recording of gross deviations from the intended flightpath.

ACKNOWLEDGEMENTS

This work was in part supported by a research grant by the CAA SRG. I would like to thank SRG staff for useful comments on drafts. This does not of course commit SRG to policy changes – although it is hoped that SRG experts will take note of factual and rational points made here. I would like to thank NATS and Eurocontrol safety experts for discussions and moral support.

REFERENCES

Apostolakis, G. E., 2004. How Useful is Quantitative Risk Assessment? Risk Analysis 24(3), 515-520.

ARIBA [ATM system safety criticality Raises Issues in Balancing Actors responsibility], 1999. WP4 Final Report: Human operators controllability of ATM safety. ARIBA/NLR/WP4/FR. http://www.aribaproject.org/rapport4/index.htm

BFU [German Federal Bureau of Aircraft Accidents Investigation], 2004. Investigation Report 'Überlingen Mid-air collision'. AX001-1-2/02. http://www.bfu-web.de/berichte/02_ax001efr.pdf

Blom, H. A. P., Bakker, G. J., Everdij, M. H. C. and van der Park, M. N. J., 2003. Collision Risk Modeling of Air Traffic. http://www.nlr.nl/public/hosted-sites/hybridge/documents/R2.9%20ECC2003-670_final.pdf

Brooker, P., 2002. 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., 2003. The Risk of Mid-Air Collision to Commercial Air Transport Aircraft receiving a Radar Advisory Service in Class F/G Airspace. Journal of the Institute of Navigation 57, 277-289.

Brooker, P., 2004a. Airborne Separation Assurance Systems: Towards a Work Programme to Prove Safety. Safety Science 42(8), 723-754.

Brooker, P., 2004b. Consistent and Up-To-Date Aviation Safety Targets. Aeronautical Journal July, 345-356.

Brooker, P., 2004c. Why the Eurocontrol Safety Regulation Commission Policy on Safety Nets and Risk Assessment is Wrong. Journal of the Institute of Navigation 57(2), 231-243.

Brooker, P., 2005a. Airborne Collision Avoidance Systems and Air Traffic Management Safety. Journal of the Institute of Navigation 58(1), 1-16.

Brooker, P., 2005b. Reducing Mid-Air Collision Risk in Controlled Airspace: Lessons from Hazardous Incidents. – to appear in Safety Science.

Brooker, P., 2005c. STCA, TCAS, Airproxes and Collision Risk. – to appear in the Journal of the Institute of Navigation.

Brooker, P., 2005d. Air Traffic Management Accident Risk Part 2: Repairing the Deficiencies of ESARR4. Cranfield University

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

Brooker, P. and Lloyd, D. E., 1978. Collision Risk and Longitudinal Separation Standards for North Atlantic Air Traffic: UK CAA DORA Communication 7801, Issue 2, 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.

DNV, 1997. Hazard Analysis of Route Separation Standards. Report to Eurocontrol. Revision 3/ September.

DNV, 2003. Safety Assessment of P-RNAV Route Spacing and Aircraft Separation. Final Report TRS 052/01 for Eurocontrol.

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

Dougherty, E. M., 1997. Is human failure a stochastic process? Reliability Engineering and System Safety 55, 209-215

Embrey, D., 2004. Qualitative and quantitative evaluation of human error in risk assessment. In Human Factors for Engineers. Eds. Sandom, C. and Harvey, R.S. IEE.

Eurocontrol SRC, 2001. Risk Assessment and Mitigation in ATM, Eurocontrol Safety Regulatory Requirement ESARR4, Edition 1.0., Eurocontrol, Brussels. http://www.eurocontrol.int/src/gallery/content/public/documents/deliverables/esarr4v1.pdf

FAA, 1991. Precision Runway Monitor Demonstration Report. DOT/FAA/RD-91/5, FAA.

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

Flight, 1922. The Air Disaster. XIV(15), 210.

Foot, P. B., 1994. A review of the results of a Trial Hazard Analysis of airspace sectors 24 and 26S, CS Report 9427, CAA, London.

Fumerton, R. and Kress, K., 2001. Causation and the Law: Preemption, Lawful Sufficiency, and Causal Sufficiency, Law and Contemporary Problems, special issue on Law and Causation in Science,, Ed. Conley, J. M. 64(4. http://www.law.duke.edu/journals/lcp/articles/lcp64dAutumn2001p83.htm

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.

Hart, H. L. A., and Honoré, A. M., 1985. Causation in the Law. 2nd edition. Oxford, Clarendon.

Hollnagel, E., 2000. Looking for errors of omission and commission or The Hunting of the Snark revisited. Reliability Engineering and System Safety 68, 135-145.

Honoré, A. M., 2001. Causation in the Law. The Stanford Encyclopedia of Philosophy, Winter 2001 Edition. Ed. Zalta, E. N. http://plato.stanford.edu/archives/win2001/entries/causation-law

Hume, D., 1777. An Inquiry Concerning Human Understanding. Ed.Selby-Bigge, L. A. Oxford University Press, 1975.

ICAO, 1956. Accident Digest Circular 54-AN/49, 95-111.

ICAO, 1976. Methodology for the Derivation of Separation Minima applied to the Spacing between Parallel Tracks in ATS Route Structures. ICAO Circular 12—AN/89/2, Second Edition.

Kirwan, B., 1994. A Guide to Practical Human Reliability Assessment. Taylor and Francis, London.

Mackie, J. L., 1980. The Cement of the Universe. A Study of Causation. Oxford, Clarendon.

Marais, K., Dulac, N. and Leveson, N., 2004. Beyond Normal Accidents and High Reliability Organizations: The Need for an Alternative Approach to Safety in Complex Systems. Engineering Systems Division Symposium, MIT, Cambridge, MA. http://sunnyday.mit.edu/papers/hro.pdf.

Mill, J. S., 1904. A System of Logic, Ratiocinative and Inductive:…8th edition, Longmans, Green, London.

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

NASA, 2005. Third Interim Report – NASA Return to Flight Task Group. http://www.returntoflight.org/assets/pdf/report-01-28-2005.pdf

Neil, M., Malcolm, B. and Shaw, R., 2003. Modelling an Air Traffic Control Environment Using Bayesian Belief Networks. 21st International System Safety Conference, Ottawa, Ontario, Canada.

Nunes, A. and Laursen, T., 2004. Identifying the factors that led to the Ueberlingen mid-air collision: implications for overall system safety. Proceedings of the 48th Annual Chapter Meeting of the Human Factors and Ergonomics Society, New Orleans, LA, USA. http://www.aviation.uiuc.edu/UnitsHFD/conference/humfac04/nuneslaur.pdf.

Perrow, C., 1984. Normal Accidents: Living with High-Risk Technologies. Basic Books, New York.

Reich, P. G., 1966. Analysis of Long-range Air Traffic Systems: Separation Standards. Journal of Navigation 19, 88-98, 169-193 and 331-347.

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

Sharpe, A. G., 1991. Application of the 5 nm radar standard separation at ranges up to 160 nm from Claxby, Debden and Pease Pottage SSRs. CAA Paper 91013. Civil Aviation Authority, London.

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

UKAB [Airprox Board], 1999 onwards – biannual. Analysis of Airprox in UK Airspace. www.ukab.org.uk.

Weick, K.E., 1976. Educational organizations as loosely-coupled systems, Administrative Science Quarterly 21(1), 1-19.

Williams, G. L. and Hepple, B. A., 1976. Foundations of the Law of Tort. Butterworths, London.

Wright, R. W., 1988. Causation, Responsibility, Risk, Probability, Naked Statistics, and Proof: Pruning the Bramble Bush by Clarifying the Concepts. Iowa Law Review, 73, 1001-1077.

Wright, R. W., 2001. Once More into the Bramble Bush: Duty, Causal Contribution and the extent of Legal Responsibility, Vanderbilt Law Review., 54, 1071-1132. http://ssrn.com/abstract=254875 (leads to pdf file)

Air traffic management accident risk, part 1: the limits of realistic modelling

Brooker, Peter

2005-06-27T13:40:21Z