Influences of communication structural complexity on operational safety in regional airspace design

Neale L. Fulton[a] **Mark Westcott**[a] **Stephen Emery**[b]

[a] *Mathematical & Information Sciences,*
Commonwealth Scientific & Industrial Research Organisation (CSIRO)
GPO Box 664, Canberra, ACT 2601, Australia
Email: neale.fulton@csiro.au

[b] *School of Civil and Environmental Engineering, University of the Witswatersrand*
Johannesburg, South Africa

DRAFT Feb 2010 v7

Abstract

As air traffic management systems have evolved and as traffic flows have increased, the communications protocols and supporting structures have become more complex. Initially, in the early 20th century, pilots flew without any inter-aircraft communication. Then area frequencies were introduced to cover large regions. More recently, extra frequencies have been introduced with a result that the larger "area" airspace volumes have become isolated in frequency from the smaller volumes that surround airports. This increasing complexity has introduced engineering failure modes into the design. Within this increasingly complex environment, regardless of the relative position and geometric aspect of an aircraft pair the pilots must be able to communicate with each other. They need to be able to rely on a high dependability (a reliability and safety concept) of the various communication links operating between the aircraft at the time of proximity. This paper is concerned with assessing the physical feasibility of inter-pilot communication when their aircraft are in, or near, radio frequency boundaries. The distinctive feature of the boundaries studied is that communication occurs on different frequencies at different points in airspace. This means that pilots in relatively close spatial proximity might not be operating on a common frequency. When combined with relatively long transaction times, this has the potential to fatally inhibit timely exchange of information critical to successful avoidance of a midair collision. We show that failure modes similar to failure "modes" of operation (e.g., mode confusion within the pilot-machine interface) identified in Flight Management Systems designs and as discussed in the aerospace literature arise for the boundary structures. The paper uses a simple but revealing model of aircraft operation within a multiple radio frequency structure to study these operational modes. The model is not intended to be complete or exhaustive; its role is to demonstrate design principles and processes that should be considered in order to achieve required levels of system design confidence. One important conclusion is that circumstances in which problems can arise are not easily predictable during flight. This means that operational experience is not necessarily a good basis on which to predicate the extrapolation of system design behaviour, as aircraft might often operate close to a failure mode without the pilots realising it and so they may erroneously conclude that these modes do not exist. The model allows an exhaustive description of the failure modes once parameters such as the aircraft velocities, the radio frequency structures and the communication transaction lengths are specified. To show how the failure modes are influenced by these parameters, the paper uses a novel form of nested plot for high-dimensional data that was developed for similar displays in a large commercial contract.

Keywords: airspace design, aircraft proximity, collective risk, design space, individual risk, mid-air collision.

1. Introduction

Air Traffic Management (ATM) systems have undergone significant changes over the past two decades. There has been considerable modernisation of ground-based ATM equipment and of the communication infrastructure (e.g., introduction of data links). This imparts new engineering demands on these systems, which are now being stressed in ways that can either cause failure modes once latent to manifest or new failure modes to emerge. Concurrently there has been an increasing use of all airspace by both new and established groups: the number of recreational sports aircraft have increased significantly; the Unmanned Aerial Vehicle (UAV) industry is reported to continue to be the most dynamic growth sector of the world aerospace industry, projected to total just over $62 billion in the next ten years (Teal, 2010); the personal jet and regional jet markets are growing strongly (Mozdzanowska, Delahaye, Hansman, Hinston, 2003); commercial and regular public transport services are increasing aircraft numbers as demand for air travel rises (Babikian, Lukachko, Waitz, 2002).

In this context of change a midair collision is one of the most dramatic accidents that can occur in aviation, and one that it is particularly important to avoid. Air transport management systems, of which ATM systems are a part, rely on various lines of defence in the form of communication links to mitigate against the occurrence of a midair collision. The dependability of the communication links that exist between aircraft when proximate to each other emerges as a critical design issue. Such dependability directly influences the assessment of the probability of a collision given the occurrence of a proximity event.

The term communication link is used here in its broadest scientific sense. It is a means of transferring information between aircraft and has characteristics of directness, mode and latency. Direct links can form between aircraft. Indirect links rely on provision of a ground-based air traffic service. While air-to-ground links are of strategic importance, air-to-air links are of tactical importance. In this paper the multiple communication links that may exist between aircraft are considered as a parallel, hierarchical, and fault tolerant system comprised of any combination of data-link, voice and see-and-avoid communications. Each means of communication augments and aids the other links in the overall objective of providing situational awareness to the pilots of the aircraft.

The most basic level of communication is ‘see-and-avoid’; one or more of the pilots observes the potential danger and takes appropriate avoiding action. However, see and avoid is more likely to be successful if a pilot knows there is another aircraft to be seen. This usually happens through a radio transmission, from either the other aircraft or from a ground-based controller but it may happen in more modern ATM systems via information exchange over a data-link. Such a situation is called ‘alerted see and avoid’. If this does not happen, the line of defence is, at best, ‘unalerted see and avoid’, where, regardless of a pilot’s vigilance, success relies on a measure of luck.

For a voice alert to be issued and received, various actions must occur. An appropriate transmission must be made; the receiving radio must be operational and on the correct frequency; the transmission must not be blocked by other transmissions; the pilot must not be distracted. In an ideal world, all these actions would happen. In practice they might not, due to human error, deficient training, equipment failure or incorrect design of the communication system. This paper focuses on the last of these possible causes.

Our research was conducted in the context of over two decades of national airspace change in Australia and elsewhere (e.g., Erzberger, 2004) that saw aircraft communications procedures become more complex. Originally, one frequency covered a large area, often including a number of airports. The changes introduced new and smaller frequency volumes around airports, called Mandatory Broadcast Zones (MBZ) or Common Traffic Advisory Frequencies (CTAF). This meant that the larger area sectors became isolated in frequency from the smaller volumes surrounding the airports. The associated radio frequency boundaries added engineering and operational complexity into the design. Within this increasingly complex environment pilots must still be able to communicate with each other if their aircraft come into proximity. This requirement applies regardless of the relative position and aspect of the aircraft pair. Pilot-to-pilot communication must be achieved with high dependability (a reliability and safety concept) at the time and for the duration of proximity. We shall show that this is not always the case.

In Australia, new voice procedures covering operations at non-controlled aerodromes were introduced throughout the decades of change. A key feature was the use of positional broadcasts to

DRAFT Feb 2010 v7

establish situational awareness and assist pilots to see and avoid potential conflicts. These changes have most recently been consolidated in CASA (2009a, 2009b). In a report on the efficacy these procedures Ambidji (2008) stated that "*the major contributors to aircraft separation risk in non-controlled airspace [relate to] non-compliance with recommended radio and circuit procedures*" (p.1). One of the report's recommendations was "[I]mplementation of a …. pilot training and education program to improve pilot compliance with CTAF procedures" (Recommendation 1d, Ambidji, 2008). Clearly, pilots should be fully trained to be aware of, and compliant with, appropriate radio procedures. However, compliance with procedures presupposes that the procedures themselves are sound and do not have inbuilt failure modes. Our contention is that the non-augmented use of radio frequency boundaries has introduced inherent deficiencies that no level of compliance can circumvent.

Taneja and Wiegmann (2001) have observed with regard to see-and-avoid communication:

> "*Why then do [mid air collisions] occur in conditions so ideal for flying? The answer to this question is obvious both to the layman as well as the learned investigator: inadequate lookout by the pilot/pilots*".

We believe the answer is not so obvious. Other factors may be at work beyond the capability of a pilot to directly control, such as inherent deficiencies in the alerting and communication systems. A graphic example on record (Australian Transport Safety Bureau occurrence number 199900420) is where a Dash-8 passenger aircraft passed laterally within 30-60 meters of a glider. This distance translates to operation within a loss-of-control regime, with inadequate response time available for the pilots.

The methodology described in this paper was used in a study (Kubu, 2004) for Broome International Airport (BIA). Over one million collision situations were tested for adequate communication response times, with 55,000 failures recorded. These observations together with other studies conducted by BIA confirmed a decision to retain ground based traffic advisory services at the airport when, in a broader national context, such services were being generally reduced or removed. This service provides the augmentation needed to mitigate the potential failure modes introduced by the frequency boundaries.

The design question that arises is:

> *What impact do communication structures (radio frequency boundaries in this case) have on the engineering design, the operational integrity, and the dependability of pilot-to-pilot communication that must exist between the proximate aircraft in order to safely manage that proximity?*

This paper is concerned with assessing the physical feasibility of inter-pilot communication when aircraft are operating in, or near, radio frequency boundaries. It examines how the structures might affect the dependability of communication between the aircraft. The distinctive feature of the structures studied is that communication can occur on different frequencies at different proximate points in airspace. This means that pilots in relatively close spatial proximity might not be operating on a common frequency. When combined with relatively long transaction times, this has the potential to fatally inhibit timely exchange of information critical to successful avoidance of a midair collision. A further point of note is that while the focus of this research is on a voice communication paradigm the same design principles will apply to the data link communication infrastructures (e.g., ADS-B) presently being considered and introduced. This is particularly so when the human response times for both transmission and reception of information are included in the transaction length determination of the pilot-to-pilot communication link.

We show that both normal operational modes and failure modes arise for radio frequency boundary structures. The modes discussed are similar to failure "modes" of operation (e.g., mode confusion

DRAFT Feb 2010 v7

within the pilot-machine interface) identified in Flight Management Systems designs and as discussed in the aerospace literature (Degani, 1996). A novel feature of this paper is the method we use to display both types of mode of the system and their dependence on a number of design and operational parameters. The visualisation used is a nested plot which shows important features of the dependence for up to six independent parameters. It was originally developed for similar displays in a commercial project (Fulton, Westcott and Emery, 2009).

While the communication problem discussed in this paper is of interest in its own right, it directly illustrates two important ideas in the design of safety critical systems.

1. Non-linearity in a system design, inclusive of discontinuities in system output or performance is a particularly tricky problem for system design. The presence of a radio frequency boundary introduces hysteresis into the system state description through the requirement to select one of two frequencies at any given position in space. The state transition diagram becomes combinatorially complex (Fulton, 2002; Fulton, Baird and Smith, 2002, 2003).

2. Kuchar (1996) (see also Yang & Kuchar, 2002) gives a general methodology for looking at the performance of alerting systems, of which the system we discuss is an example. His method is based on the construction of hazard and alert regions in the system state space. Once the system state enters the alert region, action is taken to try to avoid the state entering the hazard region.

Kuchar's methodology requires that an alert can always be issued - that is, the system can only reach the hazard region by passing through the alert region – because the system performance measures are all probabilities conditioned on an alert being issued. We show that the system investigated does not come within the proposed paradigm because it does not have this property. Two aircraft can come into hazardous proximity without an alert being issued via the radio communication system. It is important to note that while the focus of this paper is on a voice communication system the design principles also apply to systems with data-link transactions where human response times are included in the transaction time specification.

The deficiencies we identify might be apparent from a thoughtful consideration of the system. However they do not seem to have yet been generally acknowledged. In this paper we shall use a simple yet revealing model of aircraft operating within a CTAF/MBZ system to highlight the problems which can be encountered. The kinematic simplicity of the model is deliberate. The problems are associated with the communication system design, not with the aircraft. By keeping the kinematic aspects of the model as uncluttered as possible, the communication issues can be shown more clearly.

The paper uses a simple but revealing model of aircraft operation within a multiple radio frequency structure to study the operational modes. The model is not intended to be complete or exhaustive; its role is to demonstrate design principles and processes that should be considered in order to achieve required levels of system design confidence. One important conclusion is that circumstances in which problems can arise are not easily predictable during flight operations. Kinnersly and Roelen (2007) observed:

> *"-the instrumentation did not give the pilots the information they needed in this scenario, a classic failure of design to satisfactorily address Human Factors in safety related scenarios"*

DRAFT Feb 2010 v7

In the following sense the research of this paper distinguishes between heuristic approaches to procedural specification and formal engineering design approaches (e.g., Munoz et al., 2003; Dong et al., 1997). Stanton et al. (2001) observe:

> *The call has arisen with the increased realisation that system design is no longer optimised for human operation and, under some conditions, has 'overstepped the human's capability to keep track'.*

Operational experience alone is not necessarily a good basis on which to predicate the extrapolation of system design behaviour, as aircraft might often be operated close to a failure mode without the pilots realising it and so they may erroneously conclude these modes do not exist. On the basis of pilot experience other stakeholders may come to a similar conclusion. The engineering model allows a more exhaustive and complete description of the failure modes once parameters such as aircraft speeds and headings, radio frequency structures and communication transaction lengths are specified. To show how the failure modes are influenced by these parameters, the paper uses a novel form of nested plot for high-dimensional data that was developed for similar displays in a large commercial contract.

In the context of the engineering design process we therefore seek to address four questions:

1. What is the operational influence of a radio frequency boundary on pilot-to-pilot communication when aircraft are in close proximity?
2. What impact does the transaction length have on the dependability of design?
3. What can be said about the design integrity?
4. Should the design exhibit failure modes such that the pilots cannot communicate in time to avoid a collision then are these failure modes generic to both voice communication and to data link communication?

The purpose of the analysis should be to guide both policy making and engineering design processes.

In Section 2 of this paper an overview of the regulatory and operational context is given. In Section 3 a hybrid design model is described for a radio communication structure that has now become common in aviation activities. A discussion of the results for a large set of simulated situations is given in Section 4. The implications of these results are discussed in Section 5 and conclusions drawn in Section 6.

2. Regulatory and Operational context of investigation

The study of the mid-air collision event is of significant interest to airspace designers since this event can be viewed as a system constraint both from both an operational and also from an engineering perspective. An ATM system must provide for the safe transport of aircraft through airspace without invoking a collision. Second, if a mid-air collision occurs, the occurrence of such an event can be viewed as a functional failure of the airspace sub-system. To prevent such events, rules and procedures are imposed such that aircraft/pilot must comply with these when in proximity. An integral part of the design of the rules and procedures is the development of communication procedures required to exchange information on aircraft proximity and flight-path intent. The following sections provide a summary of the regulatory framework governing this problem.

2.1 International categorisation: Flight regimes and flight operations

The rules of the air and related procedures are specified within an hierarchical, multidimensional, regulatory framework that includes both the international regulations from the International Civil

DRAFT Feb 2010 v7

Aviation Organisation (ICAO) as well as regulations imposed by the ICAO States (e.g., Australia, Canada, the U.K., etc.). In ICAO (2005) Annex 2, flights are categorised according to whether flight in cloud is permitted, Instrument Flight Rules (IFR), or alternatively, and mutually exclusively, whether there is an operational need to remain in visual flight conditions, Visual Flight Rules (VFR). Second, in ICAO (2001) Annex 11, flights are also categorised by the Class(es) of Airspace in which the flight will be conducted. This class is essentially an indicator of the level of services to be provided by ground-based facilities to a flight operating under either the IFR or VFR flight category. Annex 11 refers to Controlled and Uncontrolled airspace but modern usage is to refer to the synonymous terms, for the purpose of this paper, Managed airspace (MAS) and Un-managed airspace (UMAS). These categorisations give rise to four flight regimes as shown in Table 1.

Table 1: Categorisation of flight operations – the first two levels

	VFR	**IFR**
Managed Airspace (Classes A, B, C, D &E)	Mid Level enroute Terminal airspace	High Level enroute Mid Level enroute Terminal airspace
Unmanaged Airspace (Classes F & G)	Regional enroute and terminal airspace	Regional enroute and terminal airspace

2.1.1 The ICAO State and flight operations

At the ICAO State level further categorisation of airspace operations occurs but now the focus moves more towards whether an operation is concerned with Air Transport Operations (ATO – also known as Regular Public Transport (RPT)) a category representing scheduled public services, or general aviation (inclusive of: commercial aviation operations; charter; private scheduled services; pilot training; private flying and business aviation); or sports aviation such as gliding and other recreational in-flight sports; or finally, military aviation. Categorisation schemes at this level are not uniform from one State to another and are similar to the problems of categorisation reported by Motevalli & Salmon (2004). For example, in the U.S.A. categorisation is via the FAR Parts (e.g., FAR part 121, FAR Part 135 and FAR Part 91 etc.) whereas in Australia historical categorisation in the official accident reports has been according to: High Capacity RPT, Low Capacity RPT, General Aviation, Sports Aviation or Military operations (ATSB, 2003).

Whilst it may be more common for ATO aircraft to only operate as IFR flight category in MAS, in general, there is no restriction that they always have to do so. In fact, it is the nature of flight, and part of its utility, that there may be many transitions from MAS to UMAS (and vice versa) in the one flight for either flight category. Airspace rules and procedures are required to function dependably in all situations and in all transitions.

Throughout much of continental Australia the low level airspace below Flight Level 185 (circa 18,500 feet above mean sea level) is UMAS (Class G). Here, in this airspace, ATO, GA and Sport aviation can all become proximate as aircraft approach and depart from regional airports. An example of such operations is that at Broome International Airport where ATO aircraft such as the B737 and small GA aircraft (e.g., Cessna 172 and Piper 140) share common Class G airspace on a regular basis. This use of airspace, together with the particular communication infrastructure available, may contrast markedly with both the use of other low level airspace in Europe or North America for example, or with high level enroute airspace over most continents as another example, where operations may be primarily IFR in MAS.

On occasion, but perhaps not routinely, ATO is permitted legally to operate in UMAS under both the IFR and VFR flight categories. For example an RPT flight may cancel IFR operations to then proceed by VFR flight procedure to a destination or alternate. Alternatively an RPT flight may conduct a short transit (e.g., 70-100 NM) for scheduled flights under the IFR flight category. Some

DRAFT Feb 2010 v7

companies may specifically preclude this flight regime via a prohibition in their operations manual, while others may permit it. Regardless of the flight regime or flight operation airspace rules and procedures must provide an adequate line of defence for any combination of aircraft that specify a proximity pair.

2.2 Regional and local operations

At the local (regional) level radio procedures have been prescribed for use by IFR and VFR, either collectively or independently. To facilitate such procedures a unique radio frequency (channel) is allocated to an airspace volume. When the volume is relatively large in area (~ > 100 NM^2) the radio frequency is known as the *area frequency*. Historically, this frequency serviced the surrounding region, which might sometimes include airports. In more recent times (1991 - 2005) circular cylindrical airspace volumes centred on an airport have been introduced with a separate frequency. These volumes are known as the Common Traffic Advisory Frequency (CTAF) and the Mandatory Broadcast Zone (MBZ) (AIP, 2001), the concept of which is illustrated in Fig. 1.

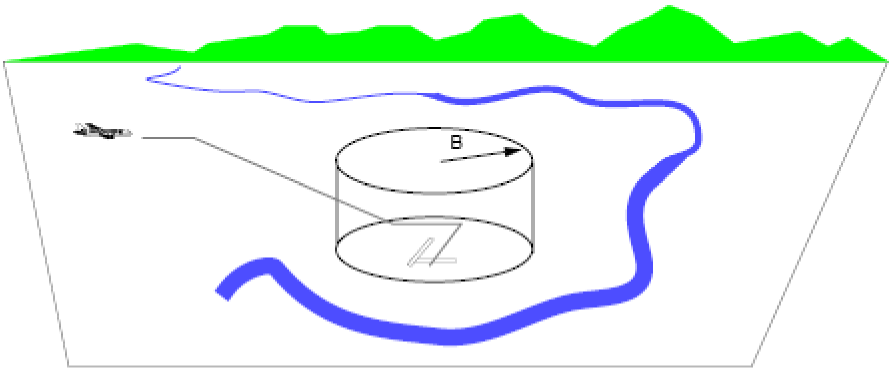

Fig. 1: CTAF/MBZ structure

For our purposes, the only difference between the CTAF and the MBZ is the radius, B, of the cylinder. The radius is typically 5 NM for the CTAF, but may range from 5 NM through circa 30 NM for the MBZ.

Radio procedures prescribe the use of an area frequency to broadcast intentions when operating outside the cylinder and on a second frequency when entering or operating within the cylinder. The rationale is to contain communication associated with circuit traffic to this second frequency, removing the need to service airport communication tasks from the area frequency.

While the intent of these structures is clear it is also obvious that communication between aircraft can degenerate when aircraft operate near to or have to transition a radio frequency boundary. This arises simply because two aircraft in close proximity can simultaneously have different frequencies selected on their radios.

Historically, the guarding of such frequency change-overs was achieved through a parallel ground-based communication function such as Directed Traffic Information (DTI). Although the implementation of this function became cost prohibitive it can now be implemented in modern technology by pilot-to-pilot data link (PPDL) or other ground-based implementations such as certified ground based controllers, or tower services. However, direct radio communication between the pilots is still a crucial part of the alerted see-and-avoid process, and such communication can still be deleteriously affected near a radio frequency boundary.

A fundamental design requirement is for pilots of proximate aircraft to be able to achieve a prescribed alert time. Fulton (2002) and Fulton, Baird and Smith (2003, 2002) investigated the reduction in alert times as a function of closing speed. When aircraft operate on the one area frequency, then most aircraft can achieve a five minute alert time if they report inbound at 30 NM. Only the fastest aircraft with the highest relative closing speed (250 - 500 Kts) fail to meet this criterion achieving instead a three minute alert when relative closing speed combinations are in the range 250 - 500 Kts.

When a radio frequency boundary is introduced then the alert times can no longer be guaranteed. Pilots may need schedule additional activities to synchronise their respective frequencies to communicate particularly if each aircraft is equipped with only one radio. This additional tasking could require that pilots should be given additional warning to switch the radio frequency early if a situation that may require proximity management is developing near the boundary, or, alternatively, that some form of boundary alert be available. Without these precautions, aircraft can come into close proximity with no prior voice warning and with insufficient time to manage the proximity successfully. The following section provides a model that explores some of the engineering complexity involved in the design of this type of system.

3. A hybrid design model for an MBZ/CTAF

In the context of this model the term "hybrid" means a model that accounts for the

1. continuous dynamics (kinematics) of the aircraft involved
2. discrete nature of the logical communication process needed to control proximity, and
3. various levels of interaction between the aircraft.

In assessing the performance of an MBZ/CTAF structure it is critical to consider the hybrid nature of the problem. The main points we emphasise are:

- There are certain combinations of kinematic and logical conditions where MBZ/CTAF structures will fail.
- ICAO (2001) Annex 11 requires continuity of service.
- The MBZ/CTAF procedures when operated under self-separation provide no warning as to when failure will occur.
- If self-separation (segregation) functions are augmented by separation functions then failure instances can be reduced. However, this is no guarantee, as demonstrated by Sioux Lookout accident (TSB-Canada, 1995).
- Pilots may use the procedures many times without failure only to find that, on the next operation, a small deviation in procedures or timing takes the system to a failure mode (see Degani (1996) on flight management systems).

3.1 The model

At time zero, an outbound aircraft, O, takes off from an airport. There is another (inbound) aircraft, I, in the area on a collision course with O. Each is initially unaware of the presence of the other. So they will collide unless they communicate in time to take appropriate evasive action. Note that I could be planning to land at the airport or just be in transit in the vicinity of the airport. We make the following assumptions throughout.

Kinematic assumptions

DRAFT Feb 2010 v7

K1. Each aircraft has a constant speed and heading, so their velocity vectors, $\mathbf{V}_O$ and $\mathbf{V}_I$, are constant.
K2. The initial position of I is such that O and I will collide if their current courses and velocities are maintained.

Communication assumptions

C1. There are two communication frequencies used; an *inner frequency* f_1 in the neighbourhood of the airport, and an *outer (area) frequency* f_2 elsewhere.
C2. Aircraft on different frequencies cannot communicate.
C3. There can be only one transmission at a time on a frequency; all other transmissions on the frequency are *blocked*.
C4. O broadcasts on f_1 as it takes off. It then broadcasts on f_2 as it exits a cylindrical region C_b of radius *b*.
C5. I broadcasts on f_1 as it enters a cylindrical region C_B of radius *B*. If it later exits C_b it will broadcast again on f_2.
C6. The transmission lengths for each aircraft, u_O and u_I, are constant but possibly different.
C7. If the aircraft communicate successfully they will avoid a collision. Successful communication means that a *complete* transmission is made by one aircraft and received by the other aircraft.

Other assumptions

A1. O and I can manoeuvre instantly to avoid a collision, so pilot and aircraft control reaction times are ignored.
A2. *B* > *b*. (this is to avoid extra complications, for purposes of exposition)

These assumptions are clearly simplistic. However the results that come from them show that there is a considerable variety and complexity of possible interactions between O and I even in this case, which is the main point of the paper.

The plane defined by $\mathbf{V}_O$ and $\mathbf{V}_I$ is called the *engagement* plane, and all planar figures are drawn in this plane. So we can define a Cartesian coordinate system in this plane, with origin at the airport, and without loss of generality we assume O flies along the *x*-axis; thus $\mathbf{V}_O = V_O(1,0)$, where V_O is the speed of O. We write V_I for the speed of I, and γ for the *speed ratio*; thus $\gamma = V_I / V_O$.

Assumption K2 implies that the *relative velocity vector* $\mathbf{V}_R = \mathbf{V}_I - \mathbf{V}_O$, which also lies in the engagement plane, must go through the origin. This is a partial explanation for the complexity of the results; the kinematics are most naturally analysed in relative velocity space while the communications rules relate to absolute space. The situation is shown in Fig2, where θ is the track intercept angle (it is 180 – the relative heading of I) and ϕ is the polar angle of $\mathbf{V}_R$. Another way to view K2 is that the position of I at $t = 0$ is on a line through the origin at angle ϕ. Its (polar) distance from the origin is written as ξ_0. The *time to collision,* t_c, is then given by $t_c = V_R.\xi_0$, where $V_R = \|\mathbf{V}_R\|$ is the relative speed of I.

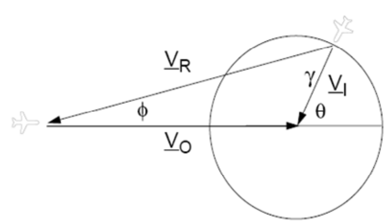

Fig2: Velocity space for proximate aircraft

3.2 Operational regimes and modes - success and failure

Because we have assumed that the two aircraft can react instantly to avoid a collision once they have communicated fully, it follows that *a collision can only occur during a transmission by one or*

DRAFT Feb 2010 v7

other aircraft. This means that the problem naturally partitions into a number of *communication regimes*. Provided the track of I does not exit C_b before meeting the track of O, these are defined as follows. Here, t_I is the time when I enters C_B and t_O is the time when O exits C_b.

Table 2: Regimes

Regime	Description	Definition
Regime 0a	I enters C_B before time $-u_I$, so finishes transmitting (on f_1) before O takes off	$t_I < -u_I$
Regime 0b	I enters C_B at time between $-u_I$ and 0, so its transmission blocks that of O at takeoff	$-u_I < t_I < 0$
Regime 1	I enters C_B while O is still transmitting. It tries to transmit (on f_1) but is blocked.	$0 \le t_I < u_O$
Regime 2	I enters C_B after time u_O but before O exits C_b, and transmits on f_1.	$u_O \le t_I < t_O$
Regime 3	I is still outside C_B when O exits C_b and transmits on f_2.	$t_I \ge t_O$

Each regime is associated with a *communication failure mode*, or *mode* for short, as follows.

Table 3: Modes

Mode	Description	Definition
Mode 0a	Collision during initial transmission by O not blocked by I's transmission i.e. during Regime 0a	$t_c \le u_O$
Mode 0b	Collision during initial transmission by O that was initially blocked by I's transmission i.e. during Regime 0b	$t_c \le t_I + u_I + u_O$
Mode 1	Collision during transmission by I that was initially blocked by O's transmission; i.e. during Regime 1.	$u_O \le t_c \le u_O + u_I$
Mode 2	Collision during transmission by I at entry to C_B; i.e. during Regime 2	$t_c \le t_I + u_I$
Mode 3	Collision during transmission by O at exit from C_b; i.e. during Regime 3	$t_O \le t_I \le t_O + u_O$

A mode 0a failure might seem contrary, since aircraft I will have transmitted on f_1 prior to O's departure. However, if O was communicating with a tower prior to takeoff, on another frequency, I's transmission would not have been heard. We are not prescribing this as a communication protocol, but just showing the consequences of the inherent mechanisms.

A mode sometimes coincides completely with a regime, particularly when I is fast. For example, in Regime 0a I can be so far inside C_B when it finishes transmitting (before O takes off), that it will always collide with O before O finishes its initial transmission (Mode 0a). A mode can also be empty; for example, if $B - b > V_R u_I$ then Mode 2 is empty because there can never be a collision within a time u_I during Regime 2. In practice, all regimes and modes are likely to be realisable.

The requirement that "the track of I does not exit C_b before meeting the track of O" always holds for θ acute. In some cases for θ obtuse it does not hold, and then we have to include the extra possibility of interacting communications when both O and I are exiting C_b.

The mathematical details of the analysis of the model are given in the Appendix. The results of the analysis are presented in the next section.

4. Results and discussion of model

There are various plots that can illustrate the results derived in the Appendix. In Fig3 we plot the track of O and a range of tracks

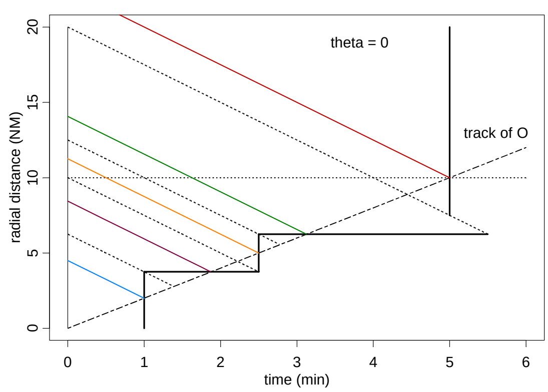

Fig3. Plot of radial distance against time for O and I, showing the Regime and Mode boundaries

DRAFT Feb 2010 v7

possible for I, specifically the Regime and Mode boundaries. The coordinates are time and the radial distance from the origin. The units are NM (distance), knots (speed) and minutes (time).

The parameter values are $v_0 = 120$, $v_I = 150$, $u_O = 1$, $u_I = 1.5$, $\theta = 0$. They have been chosen to ensure all regimes and modes are represented. This is a head-on conflict, which simplifies the formulae in the Appendix and makes all the boundaries straight lines. The dotted black lines sloping downwards are the Regime boundaries; the solid coloured lines are Mode boundaries. They come from (A.5). The solid step line is the communication boundary. The fate of a particular I, that is a particular downward sloping line defined by choice of ξ_0 (the y-intercept), depends on whether it meets the communication boundary before or after it meets the track of O.

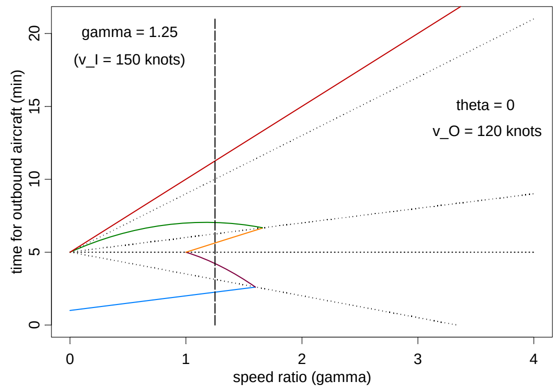

Fig4. Plot of speed ratio against 'initial position' (see text) , showing the Mode boundaries. The collision region is to the right of the solid line.

If before this, then communication occurs and the potential collision is avoided; if after, the collision occurs. Note that each mode can be either present or absent within its regime.

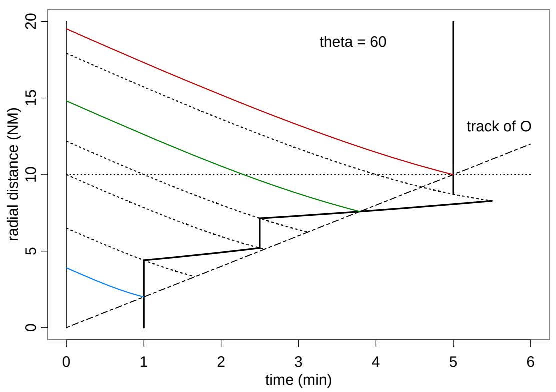

Fig5. Plot of radial distance against time for O and I, showing the Regime and Mode boundaries

The same situation is shown in Fig4 but in a plot of 'initial position' against speed ratio γ. This plot has been standardised by dividing by v_O, so the 'initial position' is actually ξ_0 / v_O, a time. Here, the safe region is to the left of the solid line while collisions occur to the right. The colours of the Regime boundaries match those in Fig3.

The dashed vertical line in Fig4 is at $\gamma = 1.25$, which is equivalent to the choice of speeds in Fig3. The switching between safety and collision as ξ_0 varies along the dotted line summarises Fig.3. Note that, for any vertical line, there is either a safe region in both Regimes 0a and 0b or in neither, and similarly for Regimes 1 and 2.

Fig4 shows the outcomes for a range of γ but it lacks the dynamic element of Fig3. Both plots are useful summaries. Each highlights the point that a very small change in circumstances can make a dramatic change in the outcome.

The next two figures show similar plots but for $\theta = 60°$. As shown in Fig.5 there is now no collision possible in Regimes 0b and 1 since a path in these regimes always reaches the

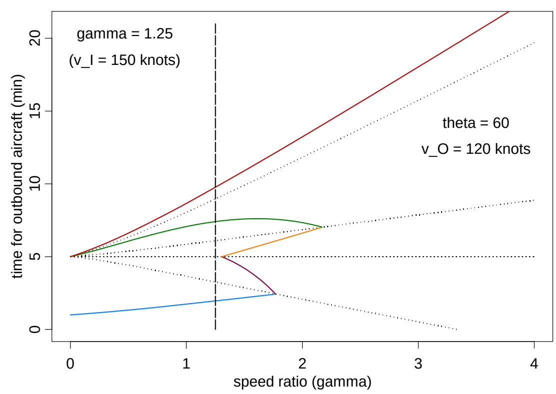

Fig6. Plot of speed ratio against 'initial position' (see text) , showing the Mode boundaries. The collision region is to the right of the solid line.

communication boundary before meeting the path of O; the corresponding modes are empty. Note also that nearly all the boundaries are no longer straight lines.

The ‘no collision’ outcome can be seen also from Fig.6, since the vertical line is now always to the left of the mode boundary lines on Regimes 0b and 1. Again, most of the boundaries are now curved, though this is not so obvious.

The angles in all of these plots are acute. When θ is obtuse, the green boundary can curve back on itself, so can be cut twice by a vertical line. This shows the added complexity that can occur for such θ. We have not dealt with such cases in any detail in this paper.

It is clear from plots like Figs. 4 and 6 that collisions become more probable as the speed ratio of the aircraft increases. The plots also suggest that collisions become less probable as θ increases. Both these conclusions accord with common sense. It is not obvious from the plots what might happen as other parameters of the model change. We could produce a large number of such plots for a range of values of each parameter, but we need to be able to view them simultaneously. We now show how this can be done.

We use a novel method for plotting high-dimensional data, which was developed for another project by the CSIRO authors.

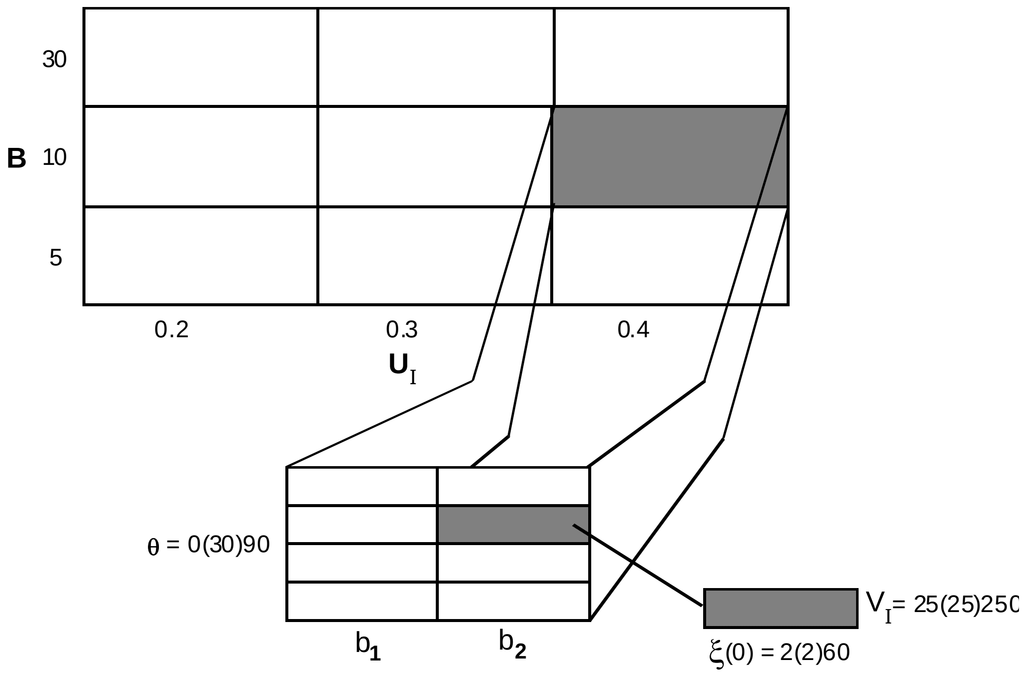

Fig.7. The template for Fig.8

The MBZ/CTAF design space in our model has eight parameters: $b, B, u_O, u_I, v_O, v_I, \theta, \xi_0$. Our plot, called a *nested plot,* can illustrate results for cases in which six of the eight parameters vary. For each individual plot, v_O and u_O are fixed. There are then three levels of nesting, as shown in Fig.7. At the top level there is a 3x3 array for the values of B and u_I. Within each of these 9 rectangles is a 4x2 array for the values of θ and b; here, $b_1 = B - 0.5, b_2 = B$. Finally, within each of these 8 rectangles is a 30x10 array for the values of ξ_0 and v_I.

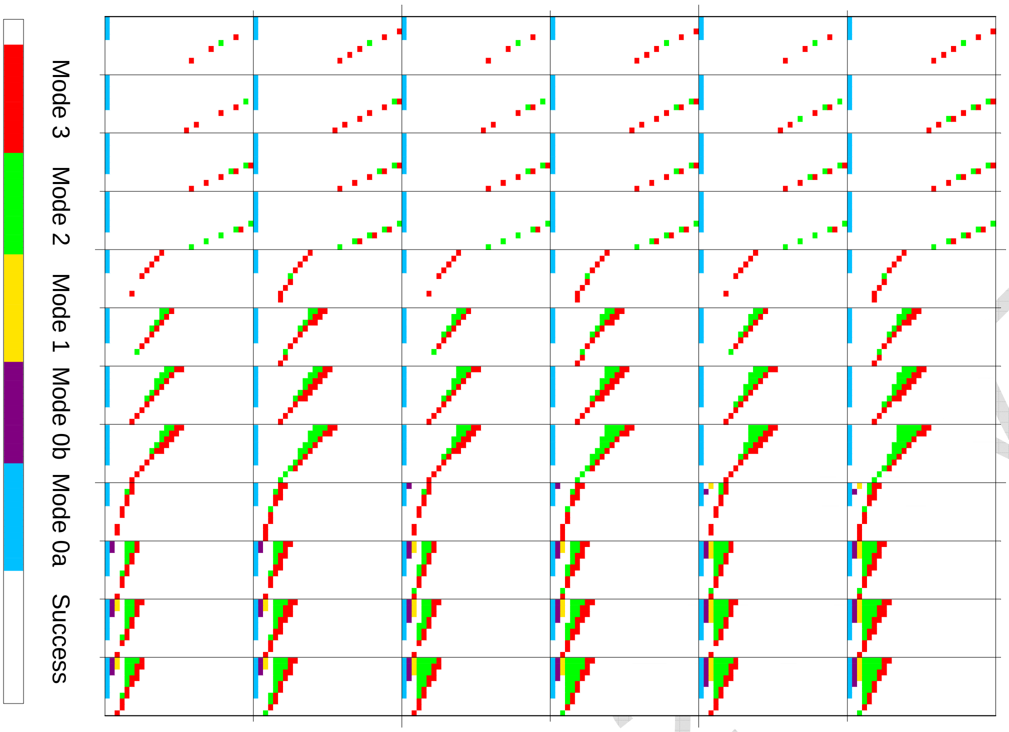

Fig. 8. Nested plot of outcomes for cases with $v_O = 120$ kts and $u_O = 0.6$ min. The template for this plot is given in Fig.7

By this means, results for a total of 21,600 cases can be shown on a single plot, colour-coded by outcome; here, the outcome is successful communication or collision, with the latter further coded by the same colours as in Figs 3-6 to show the failure mode. Note that each inner rectangle in the nested plot is essentially a discretization of Figs.4 or 6 (with the axes reversed). The number of levels of nesting, and of cases included, is partly determined by the plotting resolution, to ensure individual cases, or pixels, can be clearly distinguished.

The results of the 21,600 cases with $v_O = 240$ kts and $u_O = 0.6$ min are shown in Fig. 8. From the plot we can readily see trends in the failure patterns as the design variables change. For example: increasing B has a significant effect on the pattern whereas increasing u_I has very little effect; the results for $\theta = 0, 30$ and 60 are fairly similar but perceptibly different from those with $\theta = 90$; Mode 1 failures are rare except at the lowest value of B. We anticipated some of these findings in discussion of the earlier plots, but here we can see them clearly and simultaneously. Further, the results for $\theta = 90$ often show just a few failure cases somewhat isolated from each other. This is another instance of the points we wish to make in the paper:

- failures can be unlikely, which ensures that experience might not be a good design precedence;
- failures are hard to foresee or anticipate because the circumstances which cause them are little different from those leading to a safe outcome.

A series of such plots for different v_O and u_O gives an easily assimilated picture of the effects of variables and trend as they change. In the study for which this model was developed, we displayed the results for over 1 million cases by this means.

We mentioned in the Introduction the hazard alerting system of Kuchar, which requires that an alert region exists. Figs.3-6 are examples of a system which will not have such regions in all cases without some supplementary form of alert for pilots. For instance, a flight tracking down an arbitrary line in Fig.3 will get no effective alert before a collision if the track is in one of the failure mode regions, *and there is no indication to the pilot that the aircraft tis in such a region.*

5. Implications for design

The design of communication procedures and protocols needs to consider all aircraft with their disparate flight regimes, operational categories and performances and the manner in which they intermingle in a common airspace. Interactions of heterogeneous operational categories (e.g., Air Transport Operations (ATO) – General Aviation (GA), ATO - Gliding, GA-Gliding, etc.,) are of particular concern especially in regional airspace. For example, in Western Australia a diversity of aircraft types now operate in low level UMAS (regional) airspace (CAPA, 2002). Boeing B737-800 with 177 seats, through mid-size to middle sized regional Dash-8 aircraft (with 19 to 36 passenger seats), to small GA aircraft of 1-4 seats all sharing common airspace. Should any two of these aircraft collide then the consequence could range from zero (single pilot survives) through 42 (passengers plus crew) to 366 fatalities if all involved were to die. Design validation based on the historical accident data does not explicitly reflect the emerging potential level of collision risk. In this context avionics and air traffic control technology, have realised a useful life of ~20 to 30 years, but today such systems may be changed or replaced on a much more frequent basis. In these latter cases reliance can no longer be placed on historical data to assess the dependability of communications in managing proximity.

The main focus of future design must turn from long-term historic measures of risk to include new approaches based on an understanding the nature of traffic flow interactions, the generation of proximity within those flows and the dependability of the lines of defence provided to manage that proximity. Ultimately, this translates to setting performance requirements and designing out the latent failure modes that impair the present communication links.

Future implementation requires that the dependability of communication links between all proximate aircraft be given priority as a critical design issue that directly influences the assessment of the probability of a collision given the occurrence of a proximity event (Fulton, Westcott and Emery, 2009).

The term communication link is used here in its broadest scientific sense. It is a means of transferring information between aircraft and has characteristics of directness, mode and latency. Direct links can form between aircraft. Indirect links rely on provision of a ground-based air traffic service. While air-to-ground links are of strategic importance, air-to-air links are of tactical importance.

The physical links include: “see-and-avoid”, direct voice communication (VHF, UHF, or even HF radio), ground-based radar and data links (such as ADS-B). The physical availability of each link is a dynamic concept and multiple links may be in use simultaneously. In this situation the natural model is that of a parallel redundant heterogeneous communication network. The overall airspace design requires that both the direct and indirect links augment each other since any of the links may not be available when required, may fail, or may simply not be rapid enough to respond to an immediate proximity situation at hand

Progressively data links are being introduced to augment, but not replace, the historic communication links. Once implemented, data-links should remove much of the variation in performance now experienced due to the failure modes of the existing heterogeneous links.

DRAFT Feb 2010 v7

However, when the human latencies are added to the data link transaction times these links are shown not to be totally free of the kinds of failure modes discussed in this paper. Specifically, the dependability measures of reliability, availability, continuity-of-service, and integrity are required to be demonstrated for all short-term, on-demand inter-aircraft communications used during sustained proximity. A point to note is that the performance of the proximity management function as implemented in different parts of the world may in fact be very different simply because of the differences in the local communication infrastructures underpinning that function. Regardless of changes in infrastructure the proximity warning function must continue to function dependably as a flight progresses from one locality to another. This has not always been the case in past systems but such differences should be minimised in future implementations.

6. Conclusions

Direct voice radio communication between aircraft, particularly in the presence of radio frequency boundaries has been investigated. Five modes of failure have been identified. These modes will, a priori, be transparent to pilots and should one of these modes be encountered the system will fail catastrophically, that is, without warning.

The present analysis is based on constant velocity vectors and therefore straight line tracks. It is a first order representation that illustrates the problems without undue detail. *More realistic kinematics would change the detail but the same qualitative conclusions would emerge.* The communication policy selected is a reasonable interpretation of current operational practice. Of course, navigation errors, weather, pilot task loads and other factors all contribute to the outcome so an analysis using this model, though extensive, represents only part of the domain of possible behaviours.

The operational consequences of our results are as follows. A pilot may operate in and out of an MBZ/CTAF many times and not experiencing failure, but then a combination of circumstances (change in the values of the design variables) can occur that moves the system from a safe state to a failure mode. *And this can happen without the system providing any warning it will now fail.* There is a strong analogy between the type of failure experienced in airspace and that of software in computing; certain in-flight situations (threads) will always work, others will always fail.

Our results may partly explain the differing perceptions of different user groups within the aviation community. The performance of the MBZ/CTAF structure is highly sensitive to the closing speed between aircraft and the radius of the structure. So two slow gliders might claim the communication protocol always works; two high performance RPT aircraft operating at the legal maximum of 250 KTs (below 10,000 FT in Australia) might claim frequent difficulties with the protocol while two intermediate speed GA aircraft might claim that the protocol works most of the time.

The clear conclusion is that there will be situations in which unaided self separation (segregation) by voice link will, at times, fail in both the MBZ and the CTAF structure. Self separation needs to be augmented by some other form of communication that, jointly with voice, creates a fault tolerant system. This will typically be by a ground-based agency but with modern mobile communications such augmentation may be implemented by an air-to-air function such as and pilot alerting via aircraft-to-aircraft exchange of kinematic information by data-link.

Historically, the practical measurement of flightpath activity in unmanaged airspace has been a difficult problem due to the lack of navigation precision, systematic real-time reporting and recording of aircraft position. In addition sporadic and often low frequency of activity has meant that accumulated operational experience cannot populate the required combinatorially large state-space to be investigated even after decades of use. In terms of airspace design and analysis our

DRAFT Feb 2010 v7

model provides a systematic mathematical basis by which the operational state-space can now be specified, populated and explored. It allows the adequacy and feasibility of various communication protocols and procedures to be tested and compared both by analysis and simulation.

In terms of operational test the practical difficulties of measuring in real-time, aircraft position, flightpath propagation and interaction will be largely overcome, as data-link communication between aircraft becomes an operational reality. In this situation the mathematical modelling presented will permit a more accurate dynamic forecasting of system performance and identification of limitations. In particular, data-link reporting of position will also provide a foundational specification on which a more accurate monitoring and prediction of the actual performance of still essential aircraft-to-aircraft voice communication can be based.

References

AIP (2004) Aeronautical Information Publication, Civil Aviation Safety Authority, Canberra, ACT, Australia.

ATSB (2003), Aviation Safety Indicators 2002, A Report on safety indicators relating to Australian aviation, Australian Transport Safety Bureau, Department of Transport and Regional Services, Canberra, ACT, Australia, November.

Ambidji (2008), CTAF versus CTAF(R) Study – Final Report, 30 September 2008, The Ambidji Group Pty Ltd, Victoria, Australia. Office of Airspace Regulation File Ref: EF08/38.

Babikian, R., Lukachko, S. P., Waitz, I.A., (2002), Historical Fuel Efficiency Characteristics of Regional Aircraft from Technological, Operational and Cost Perspectives, Journal of Air Transport Management, 8 (6) 389-400, Nov.

CAPA (2002), Centre for Asia Pacific Aviation & Tourism Futures International, Review and Assessment of the Effectiveness of Air Services in Western Australia - Technical Report For Department for Planning and Infrastructure, Western Australia.

CASA (2009a), Draft CAAP 166-1(0): Operations in the vicinity of non-towered (non-controlled) aerodromes Civil Aviation safety Authority, Canberra, ACT, Australia, September.

CASA (2009b), Draft CAAP 166-2(0): Pilots' responsibility in collision avoidance in the vicinity of non-towered (non-controlled) aerodromes by 'see and avoid', Civil Aviation safety Authority, Canberra, ACT, Australia, September.

Degani, A., (1996), Modelling Human-Machine Systems: On Modes, Error, and Patterns of Interaction, PhD Dissertation, Georgia Institute of Technology, December.

Dong, J., Fulton, N., Zucconi, L., Colton, J., (1997), Formalising Process Scheduling Requirements for an Aircraft Operational Flight Program. In Proc. IEEE Computer Society First International Conference on Formal Engineering Methods, Hiroshima, Japan, IEEE, November 12-14.

Erzberger, H., (2004), Transforming the NAS: The Next Generation Air Traffic Control System, In Proc. 24th International Congress of the Aeronautical Sciences (ICAS), Yokohama, Japan, August 30.

FAA (2007), Federal Aviation Regulations (FAR), Federal Aviation Administration, U.S. Department of Transportation, Washington, D.C.

Fulton, N.L., (2002), Regional Airspace Design: A Structured Systems Engineering Approach, Doctoral Dissertation, University of New South Wales, Canberra, ACT., Australia.

Fulton, N.L., Baird, J., Smith, W., (2002), Modelling Airspace Design from a Systems Engineering Perspective, In Proc. 9th AIAA/ISSMO Symposium on Multi-disciplinary Analysis and Optimisation, Atlanta, 4-6 September.

Fulton, N.L., Baird, J., Smith, W., (2003), A Compromise Decision Support Problem Approach to Airspace Design, In Proc. 10th Australian International Aerospace Congress/ 14th National Space Engineering Symposium, Brisbane, 29 July- 1 August.

Fulton, N.L., Westcott, M., Emery, S., (2009), Decision support for risk assessment of mid-air collisions via population based measures, Transportation Research Part A, 43, 150-169.

Fulton, N.L., Westcott, M., Emery, S, (2009), *Simulating radio communication complexity in regional airspace design,* In Anderssen, R.S., R.D. Braddock and L.T.H. Newham (eds) 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand

DRAFT Feb 2010 v7

and International Association for Mathematics and Computers in Simulation, July 2009, pp. 397-403. ISBN: 978-0-9758400-7-8.

ICAO (2001), Air Traffic Services, Annex 11 To The Convention on International Civil Aviation, International Civil Aviation Organisation, Thirteenth Edition, Montreal, Canada, July.

ICAO (2005), Rules of the Air, Annex 2 To The Convention on International Civil Aviation, International Civil Aviation Organisation, Tenth Edition, Montreal, Canada, July.

Kinnersley, S., Roelen, A., (2007), "The contribution of design to accidents", Safety Science 45 (2007) 31-60.

Kubu (2004), Design Aeronautical Study for Broome International Airport Terminal Airspace, Kubu Australia Pty Ltd, Perth, WA, July.

Kuchar, J.K., (1996), Methodology for Alerting-System Performance Evaluation, Journal of Guidance, Control and Dynamics, Vol. 19, No. 2, March-April.

Motevalli, V., Salmon, C., (2004), Developing Greater Flexibility and Resolution in Aviation Accident Analyses, In Proc. ITS Transportation Safety and Security Conference, Miami, FL, March 24-25.

Mozdzanowska, A., Delahaye, D., Hansman, R.J., Histon. J., (2003), Emergence of Regional Jets and the Implications on Air Traffic Management, In Proc. 2003 Air Traffic Management Conference, Budapest, Hungary, June.

Munoz, C., Carreno, V., Dowek, G., Butler, R., (2003), Formal Verification of conflict detection algorithms, Int. J. Softw Tools Technol Transfer 4:371-380, DOI 10.1007/s10009-002-0084-3.

Stanton, N.A., Chambers, P.R.G., Piggott, J., (2001), Situational Awareness and Safety, Safety Science 39, 189-204.

Taneja, N., Wiegmann, D.A., (2001), Analysis of Mid-Air Collision in Civil Aviation, In Proc. 45th Annual Meeting of the Human Factors and Ergonomics Society, Human factors & Ergonomics Society, Santa Monica.

Teal (2010), World Unmanned Aerial Vehicle Systems: 2010 Market Survey and Forecast, Teal Group Corporation, Fairfax, VA 22030, USA.

TSB-Canada (1995), AVIATION REPORTS - 1995 - A95H0008 Mid-Air Collision Between Bearskin Airlines Fairchild Metro 23 C-GYYB and Air Sandy Inc. Registration PA-31 Navajo C-GYPZ Sioux Lookout, Ontario 12 nm NW 01 May 1995, The Transportation Safety Board of Canada (TSB)

Yang, L.C, Kuchar, J.K., (2002), Performance Metric Alerting: A New Design Approach for Complex Alerting Problems, IEEE Transactions on Systems, Man and Cybernetics—Part A: Systems and Humans, Vol. 32, No. 1, January.

Appendix. Mathematical analysis

Nomenclature

$\underline{V}_R$	the relative velocity between the inbound and outbound aircraft
$\underline{V}_I$, V_I	the velocity, speed of the inbound aircraft
$\underline{V}_O$, V_O	the velocity, speed of the outbound aircraft
γ	the speed ratio normalised to the outbound aircraft's speed ($\lvert\underline{V}_I\rvert / \lvert\underline{V}_O\rvert$)
θ	track intercept angle ($\theta = \pi - \psi$)
ψ	relative heading angle ($\cos(\psi) = \underline{V}_I . \underline{V}_O$)
ϕ	the intercept angle the relative velocity vector makes with the outbound track
$\xi(0)$	the initial relative range of the inbound aircraft at time t = 0.
B	the inbound aircraft's estimate of the MBZ/CTAF boundary radius
b	the outbound aircraft's estimate of the MBZ/CTAF boundary radius

DRAFT Feb 2010 v7

t_I the absolute time at which the inbound aircraft reaches the outer boundary

t_0 the absolute time at which the outbound aircraft reaches the inner boundary

U_I the communication transaction time interval for the inbound aircraft

U_0 the communication transaction time interval for the outbound aircraft

We need to determine the boundaries of the regimes and modes as functions of the parameters of the model. These are defined by particular values of t_c, t_I or $t_c - t_I$. The latter two can be turned into expressions for t_c alone using the basic relationship

$$B^2 = (v_O t_c)^2 + \{v_I (t_c - t_I)\}^2 - 2 v_O t_c v_I (t_c - t_I) \cos(\pi - \theta) \tag{A.1}$$

This comes from the cosine rule applied to the triangle OCI_B in Fig.A1.

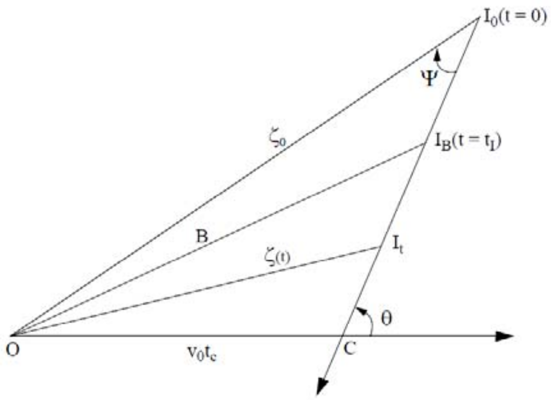

Fig. A.1 The engagement triangle

The Figure is drawn for a case where I is outside C_B at $t = 0$ (so $t_I > 0$); the changes for other cases are obvious. Here, I_0 is the position of I at $t = 0$; I_B is the position when I enters C_B ($t = t_I$); I_t is the position at an arbitrary time *t*; C is the collision point.

If $t_c - t_I$ is known, equal to τ say, we can solve (A.1) directly for t_c, getting

$$t_c = \frac{-v_O v_I \tau \cos\theta \pm \sqrt{(v_O v_I \tau \cos\theta)^2 + B^2 - (v_I \tau)^2}}{v_O{}^2}$$

$$= -\gamma\tau\cos\theta \pm \sqrt{t_B{}^2 - (\gamma\tau\sin\theta)^2} \tag{A.2}$$

where $t_B \equiv B/v_O$ is the time it would take the outbound aircraft to reach the boundary *B*.

Note that if θ is acute, so $\cos\theta$ is nonnegative, only the positive square root in (A.2) gives a physically sensible solution. However if θ is obtuse, then both the solutions in (A.2) may occur.

If t_I is known, equal to τ say, we rearrange (A.1) to get

$$t_c{}^2\left(v_O{}^2+v_I{}^2+2v_Ov_I\cos\theta\right)-2t_cv_I\tau\left(v_I+v_O\cos\theta\right)-\left(B^2-\{v_I\tau\}^2\right)=0 \qquad \text{(A.3)}$$

From Fig.2, we see that the coefficient of $t_c{}^2$ in (A.3) is V_R^2. Write $V_R^2=v_0^2\,R(\gamma)^2$. Then

$$t_c=\frac{v_I\tau\left(v_I+v_O\cos\theta\right)+\sqrt{\left(v_I\tau\left\{v_I+v_O\cos\theta\right\}\right)^2+\left\{B^2-\left(v_I\tau\right)^2\right\}v_0^2\,R(\gamma)^2}}{v_O{}^2R(\gamma)^2}$$

$$=\frac{\gamma\tau(\gamma+\cos\theta)+\sqrt{\{t_BR(\gamma)\}^2-(\gamma\tau\sin\theta)^2}}{R(\gamma)^2} \qquad \text{(A.4)}$$

In this case, only the positive square root is relevant.

These formulae show that v_O, v_I and B occur only in the combinations γ and t_B.

The most useful variable for plotting the boundaries is ξ_0, which we noted earlier is related to the collision time t_c by $\xi_0=V_R\,t_c=v_0\,R(\gamma)\,t_c$. To get the boundary values we either use a value of t_c directly or substitute the appropriate value of τ into (A.2) or (A.4). For example, the boundary between Regimes 0a and 0b comes from setting $\tau=-u_I$ in (A.4).

We use two types of plot in the displays of Section 3.3. The first plots the tracks of O and I as radial distance $\xi(t)$ versus time *t*. For O, the relationship is simply $\xi(t)=v_O\,t$. For I, we have

$$\xi(t)^2=\xi_0^2+v_I^2\,t^2-2\xi_0v_I\,t\cos\psi \qquad \text{(A.5)}$$

from the cosine rule on triangle OCI_t in Fig.A1. The angle ψ is determined by the sine rule in triangle OCI_0 in Fig.A1, which gives

$$\sin\psi=v_0\,t_c\sin\theta/\xi_0=v_0\sin\theta/V_R \qquad \text{(A.6)}$$

The regime and mode boundaries are given by (A.5) with the appropriate values of ξ_0 as determined above.

The second plot shows how ξ_0 varies as a function of the speed ratio γ. The plot is actually of the normalized variable ξ_0/v_0 which has the dimensions of time. The formulae (A.2) and (A.4) are directly applicable here, with appropriate choices for τ.

DRAFT Feb 2010 v7