Author Preprint Copy. AIAA Aviation Forum. 17th Aviation Technology, Integration, and Operations Conference. June 2017.
Final version appears on AIAA Aerospace Research Central. DOI: 10.2514/6.2017-3275

Making a Risk Informed Safety Case for Small Unmanned Aircraft System Operations

Reece Clothier*
Boeing Research & Technology – Australia, Brisbane, QLD 4001, Australia
Ewen Denney†, and Ganesh Pai‡
SGT Inc., NASA Ames Research Center, Moffett Field, CA 94035, USA

This paper describes an approach to create a Risk Informed Safety Case (RISC) towards facilitating safe, cost-effective operations with small unmanned aircraft systems (sUAS). The core constituents of a RISC are *i*) barrier models of safety, which underpin the development of a comprehensive collection of safety measures so as to be commensurate with the safety risk posed, and *ii*) structured arguments, which provide assurance of safety in operations through explicit rationale and the appropriate evidence. We also identify key safety-related assurance concerns that are relevant for improving confidence in a RISC, and, in turn, in operational safety. Then, we present a tiered framework to structure the corresponding assurance arguments. This work, which has been motivated, in part, by an absence of the relevant aviation regulations and applicable performance standards, leverages our successful, collective prior experience in creating RISCs for real sUAS operations. We draw upon this background to provide illustrative examples of our approach.

I. Introduction

TYPICALLY, access to the National Airspace System (NAS), e.g., in the U.S., Europe, and Australia, is governed by regulations and operational restrictions so that an acceptable level of safety can be established and maintained for different airspace users. Civil aviation uses a combination of highly prescriptive *normative regulations*, which mandate concrete product requirements and compliance processes, along with *performance-based regulations* that specify so-called *Minimum Operating Performance Standards* (MOPS). Due to the marked difference between the kinds of operations undertaken with conventionally piloted aircraft (CPA) versus those intended with remotely piloted aircraft systems (RPAS)—also known as unmanned aircraft systems (UAS)—applying the existing set of regulations can result in requirements and operating constraints that impede their cost-effective use.

Consequently, various national and international efforts are underway to not only adapt and tailor existing regulations, but also to develop new sets of prescriptive regulations and performance standards,[1,2] especially in the context of small UAS (sUAS). Although, rapid changes in UAS technology together with the variety in both the concepts of operation and the air-vehicle systems involved, makes it challenging for regulatory and standardization efforts to keep pace. In response, a number of *risk-based* approaches have been proposed to address various aspects of UAS safety, including:

a) characterizing safety risk in terms of the potential harm (or damage), and deriving a so-called *Target Level of Safety* (TLOS),[3,4]
b) developing safety objectives that are commensurate with the level of risk posed,[5,6]
c) deriving overall safety performance requirements,[4] as well as requirements on performance and interoperability of lower-level (flight critical) systems, e.g., avionics,[7]
d) risk modeling, and identifying the safety measures required for risk reduction,[8,9] and
e) providing assurance of airworthiness,[10] safety contribution of specific systems to be used during sUAS operations,[11] and overall system safety.[12]

The latter two items (d and e above) represent prior work by the authors that has used *barrier models*, embodied by *Bow Tie Diagrams* (BTDs), as the basis for safety risk modeling. In particular, Clothier et al., have used BTDs

*Principal Researcher, Autonomous Systems, Brisbane Technology Centre. AIAA Member.
†Senior Computer Scientist, Intelligent Systems Division. AIAA Member.
‡Research Engineer, Intelligent Systems Division. AIAA Member.

not only to assess and manage the safety risk posed during specific UAS operations, but also to underpin the safety measures put forth in the associated *safety cases*,[8,9] i.e., the risk management artifacts required to obtain regulatory approval to undertake certain kinds of UAS operations in civil airspace.

Along similar lines, Denney and Pai have used BTDs, first, to provide the risk basis for deriving airworthiness requirements for an unmanned rotorcraft system,[10] and, subsequently, in the safety case for *Beyond Visual Line-of-Sight* (BVLOS) sUAS operations.[11,12] Additionally, their work augments the risk basis with safety rationale captured using *structured arguments*, i.e., a chain of reasoning that explicitly links a series of (safety) assurance claims to concrete, heterogeneous items of the pertinent evidence. This serves as the mechanism for providing assurance that *i*) systems required for sUAS operational safety—such as those implementing a ground-based detect and avoid (GBDAA) capability—will function as intended, thereby contributing to risk reduction, and *ii*) the overall operations can be safely conducted, i.e., at an acceptable level of safety risk.

In this paper, we reconcile and further extend our prior work, elaborating our approach to develop a *Risk Informed Safety Case* (RISC)[13,14] as applied to the context of safety assurance of sUAS operations. Effectively, the combination of barrier models (that provide a risk basis), and structured arguments (capturing safety rationale) constitutes a RISC. It presents the comprehensive collection of measures taken to eliminate, reduce, or control the safety risk posed. Furthermore, it substantiates safety claims through evidence and justification, so that there is assurance that the system as designed and operated is adequate, appropriate, and effective. The main foci (and also the main contributions) of this paper are *a*) identifying key safety-related assurance concerns to be addressed by a RISC, and *b*) the development of a tiered framework for reasoning about those concerns, which leverages structured arguments to provide regulators and various stakeholders with justified confidence in operational safety.

Our paper is organized as follows: first, in Section II, we give a background on sUAS safety regulation and the broad safety risk management frameworks under which operations can be approved. Next, in Section III, we summarize how sUAS concepts of operation (CONOPS) have been generally classified, to provide a starting point for (operational) safety risk modeling, and for specifying the overall safety objectives. We also give examples of real CONOPS to concretize this classification. Based on these examples, in Section IV, we motivate the need for a risk-based safety management approach, highlighting how the nature of safety risk differs for common hazards, and can give rise to different safety measures. Then, we describe the use of BTDs, first to characterize and evaluate safety risk, and subsequently to develop the safety measures required to manage that risk. Section V presents presents key assurance concerns, and a tiered assurance framework that leverages structured argumentation to capture and convey safety rationale. The latter pertains to how the identified safety measures meet the relevant (safety and assurance) requirements, and how they contribute to risk reduction, both individually and collectively. We also elaborate how we represent the applicable arguments. Throughout the paper, we exemplify our approach using fragments of the risk models and assurance arguments as applied in a real RISC. Section VI concludes the paper. summarizing our observations and presenting some avenues for future research.

II. sUAS Safety Risk Management Background

A. Safety Regulation and Safety Cases

The primary purpose of aviation safety regulations is to ensure that the safety risks associated with aircraft operations are managed to acceptable levels. Assurance in the management of safety risks is achieved through the promulgation of procedures, standards, and licensing requirements relating to *i*) airspace integration, e.g., remote pilot training and licensing, aircraft equipage, rules of the air, etc., and *ii*) airworthiness, e.g., standards pertaining to sUAS design, manufacture, and maintenance. Currently, the risks associated with sUAS operations are primarily being managed through regulations that levy operational restrictions.[a] Additionally, any application for exemption, or permission to operate outside of the standard operating conditions defined in the regulations, requires supplying a comprehensive safety justification and/or establishing a safety case.

What is considered to be an (aviation) safety case varies based upon the aviation regulatory authority, application, or guidance document referenced.[15–19] Each of these also supplies its own context-, or application-specific interpretation of the exact nature and purpose of a safety case, together with the expected content and presentation format. Common to all are the requirements to explicitly state the overall safety objectives and requirements, and to supply the substantiating evidence. For this paper, we provide the following definition for a safety case, which we believe is largely compatible with different regulations and guidelines applicable for sUAS operations: *A safety case is a safety risk management artifact that is comprehensible, defensible, and that provides a demonstrable, and valid justification*

[a]In the U.S., the federal aviation regulations (FARs) outlined in 14 CFR, Part 107; see https://go.usa.gov/xNCQQ/

of the safety of a system and its operations, for a given application, in a defined operating environment.

Furthermore, we consider a Risk Informed Safety Case (RISC) as a safety case that includes a risk basis and assessment based on BTDs (see Sections IV.C, and IV.D), along with safety rationale captured using *structured arguments* (see Section V.B). Additionally, this artifact supplies *a*) details about the (baseline) system and environment, including existing procedures, operations, roles and responsibilities; *b*) any intended changes to the system, e.g., the introduction of new technology, equipment and procedures; *c*) sUAS capabilities, and airworthiness information; *d*) hazard and risk analyses (of the proposed changes) including details of the assumptions made, the criteria for categorizing hazards, the levels of initial and residual risk, hazard mitigations, risk treatment and hazard tracking; and *e*) details of safety risk management planning.

The aim of such a RISC for an sUAS operation is to provide assurance to the regulators of safety in operations, formulated as safety claims/objectives. A similar notion has recently been advocated by the National Aeronautics and Space Administration (NASA) in the context of system safety guidance.[13] Related to this, *objective hierarchies*,[20] seek to reformulate assurance standards in terms of objectives, or claims, that can be seen as providing the skeleton of an argument.

B. Safety Objectives

Safety objectives can be specified in terms of the acceptable management of risks. What the latter constitutes, depends on the safety decision making framework adopted by the applicable regulatory authority. For instance, the *As Low As Reasonably Practicable* (ALARP) framework has been adopted by the national aviation authorities of the U.K. and Australia.[17,21] A compatible alternative is the *So Far As Is Reasonably Practicable* (SFAIRP) framework used in defense aviation in Australia.[22]

In the U.S., the decision making framework for aviation[b] involves demonstrating compliance to the applicable federal aviation regulations (FARs), the requirements set forth in the relevant rulemaking documents,[15] and that a TLOS can be met (where required/possible). In the case of sUAS operations conducted in civil airspace by NASA, it must additionally be shown that the relevant NASA Procedural Requirements (NPRs) have been met, and where applicable/required, the *As Safe As Reasonably Practicable* (ASARP) framework is used.[13,14] In principle, this is similar to the ALARP framework in that a decision to accept a claim of acceptable safety is made when it is shown both that the safety performance achieved is tolerable and that any incremental safety performance improvement would incur an unacceptable cost in other competing measures, including performance or resources.

These frameworks specify a process for risk treatment (mitigation) and the evaluation of residual risks. It is important to note that showing that residual risk levels meet specified risk criteria may not be sufficient for a demonstration of acceptable management of risk, and consequently, the satisfaction of the safety objective: for example, the requirement to implement a *hierarchy of controls* as part of the SFAIRP framework. Thus, the satisfaction of a safety objective may warrant the substantiation of multiple claims. The interpretation of a particular safety decision making framework can depend on the legal jurisdiction in which it is applied.[23] Finally, it is not sufficient to only show that a safety objective has been met, it must also be shown how that safety objective can be maintained across the entire operational lifetime.

The safety risk management process is essential to the generation of the evidence and arguments needed to substantiate safety claims[17] irrespective of the particular safety decision making framework adopted. The application of BTDs as a specific risk modeling technique and how it can be used to support the development of a RISC is one focus of this paper.

III. sUAS Concepts of Operation

A. Operations Classification

1. *Within the United States*

In the United States, six types of access profiles[24] have been used to broadly characterize[c] how UAS can access and operate in the NAS:

Visual Line-of-Sight (VLOS): Operations can occur in Class D, E, and G airspaces, and use ground-based, sea-based, or airborne visual observers (VOs) to provide airspace situational awareness. Constraints exist on the

[b] In the Nuclear domain and for radiation/ionization safety management, the *As Low As (is) Reasonably Achievable* (ALARA) framework, which is similar to the ALARP framework, is used as per 10 CFR 20.1003. See https://go.usa.gov/xNCzS

[c] Mainly in the context of defense operations; although, with some modifications, they are also applicable and valid for civil operations.

VOs, and communication with air traffic control (ATC) may/may not be needed.

Terminal Area: Operations occur within a defined and bounded volume of Class C, D, E, and G airspace near a terminal area. Again, ATC communications may/may not be needed.

Operating Areas: Primarily concerns access to, and flight within, special use airspace including military operating areas (MOAs), restricted areas, airspace with temporary flight restrictions, etc. A variation on this access profile, which is appropriate for conducting civil sUAS operations, involves a bounded volume of airspace that is not restricted or special use airspace.

Lateral Transit (Corridor): Operations can occur in airspace classes A, C, D, E and G. These will involve a *transition corridor*, i.e., a pre-defined airspace volume, typically in Class E airspace, that laterally connects two or more defined airspace volumes, e.g., those relevant for *terminal area* and *operating area* type of operations as described earlier.

Vertical Transit (Cylinder): Involves a spiral climb/decent, where access to higher altitudes is required, e.g., Class A airspace, from within a controlled airspace, i.e., Classes C, D, and E, including terminal areas.

Dynamic: Operations represent the ultimate goal of routine access to the NAS, in much the same way as CPA, e.g., by filing a flight plan, and integrating into the NAS.

Except for the VLOS access profile, all others involve either BVLOS and/or so-called *Beyond Radio Line-of-Sight* (BRLOS) operations. The exact safety considerations required to enable these types of access profiles are not in scope for this paper. However, in brief, to operate under each access profile there is a progressive increase in the capabilities required from the UAS. Moreover, there are additional requirements on equipment, crew, procedures and the safety measures required. For instance, *terminal area* and *lateral transit* type operations may require UAS with a higher degree of airworthiness, together with surveillance capabilities (beyond those available through the existing air traffic management system) implemented either onboard the air vehicle, or located on the ground.

Here, the key observation is that the relevant safety measures required, are to be determined based on the type and nature of the safety risk posed. For instance, *corridor* type operations that occur over populated areas, or in an airspace shared with other aircraft would require assurance of an airworthy UAS, as opposed operations conducted under, say, the *operating areas* access profile in an airspace located over an unpopulated area. We note that existing capabilities of sUAS, especially those related to size, weight, power, and endurance may preclude, operations in certain access profiles, e.g., *vertical transit* type operations.

2. *Outside the United States*

The European approach to integrate RPAS/UAS operations into their NAS[25] classifies the different CONOPS and types of operation based, mainly, on

Operating Altitude: Operations are classified as:

- Very Low Level (VLL), from the surface to $500\,\mathrm{ft.}$ above ground level (AGL). These have been further classified on the basis of flight range, i.e., whether operations occur within VLOS or BVLOS.
- so-called instrument flight rules/visual flight rules (IFR/VFR) operations, from $500\,\mathrm{ft.}$ to FL600. Here, the prevailing airspace requirements applicable for CPA are expected to be retained for UAS, and with additional MOPS being levied.
- Very High Level (VHL) operations, above FL600. Here, the transit to and from this altitude poses the primary safety concern.

Traffic Class: Seven traffic classes are considered:

- Classes I, II, III, and IV are applicable during VLL operations and concern the nature of how sUAS will be operated. For example, whether they are undertake free flight, follow an air route, etc.
- Classes V, and VI occur during IFR/VFR operations, and are differentiated based on such constraints as being able to meet air route network requirements, following standard arrival/departure procedures, etc.
- Class VII traffic occurs solely during VHL operations.

The *Joint Authorities on Rulemaking for Unmanned Systems* (JARUS) consortium[d], comprising members of national aviation authorities from over fifty one nations, across all continents, has adopted the European classification to characterize sUAS CONOPS.

Again, the existing capabilities of sUAS may only facilitate certain CONOPS, e.g., VLL, and possibly IFR/VFR type operations, with the latter limited by the maximum operating altitude for the air vehicles involved. Based upon this, the admissible CONOPS for sUAS can be further refined into operational scenarios that permute over different classes of ground population density,[5] as well as airspace classes.[25]

[d]See http://jarus-rpas.org/

B. Illustrative Examples

We now give examples of real sUAS operations to exemplify the preceding CONOPS classification. Subsequently in the paper, we will also refer to these examples to illustrate our approach.

1. *The UTM Flight Trials (UFT) Example — Operations Within a Bounded Operating Volume*

In light of the growing proliferation of sUAS, NASA is developing the UAS Traffic Management (UTM) system, as a novel, highly automated, and increasingly autonomous low-altitude air traffic management service, with a view to safely integrating sUAS operations into the NAS.[26] The proof of concept involves using a fleet of sUAS to conduct a series of increasingly complex flight trials. Those, in turn, are associated with classes of CONOPS that are designed to reflect the kind of sUAS operations that that UTM will eventually support, e.g., search and rescue, disaster response, precision agriculture, package delivery, etc.

One such class of CONOPS involves a combination of VLOS and BVLOS operations with both fixed-wing and rotary-wing sUAS, within a confined volume of Class G airspace—an *operating range* (OR)—at a maximum operating altitude of 700 ft. AGL. These operations can be considered as representative of the *operating areas* type of access profile.[24] Equivalently, it is also a combination of VLL and VFR operations, involving Class III type of air traffic, i.e., following specific routes.[25]

As a matter of policy,[27] NASA requires operational approval[e] to conduct these flight trials, the process for which requires providing comprehensive safety justification, submitted in the form of a safety case.[15]

2. *The Queensland Gas Company (QGC) Example — Operations in Non-segregated Airspace*

A rapidly emerging commercial application space for sUAS is in the cost-effective long range inspection of road and railways, pipelines, power lines and other geographically distributed infrastructure. A real world example, typical of such commercial applications for sUAS, is Queensland Gas Company's use of an Insitu Pacific Ltd. ScanEagle RPAS to inspect gas wells, pipelines and processing facilities in Western Queensland, Australia.[28] This particular CONOPS requires BVLOS operations in non-segregated Class G airspace at altitudes from 1200 ft. to 1500 ft., with the operations occurring over rural and sparsely populated areas.

IV. Modeling sUAS Operational Risk

A. sUAS Hazards

The *primary* safety hazards for all sUAS operations[29] are:

i) a collision between the unmanned aircraft (UA) and a CPA, whether or not the latter is airborne, and

ii) the UA, or its components, impacting people or structures situated on the ground.

The realization of these hazards can lead to *secondary* hazards such as uncontrolled fires, falling debris, and release of contaminants, which can also pose a risk to people and property.

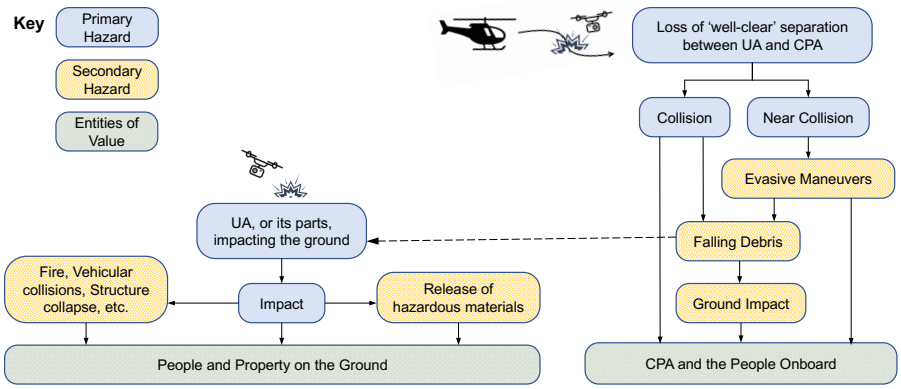


Figure 1. Primary and secondary safety hazards associated with UAS operations.

Fig. 1 illustrates the relationships between the primary and secondary hazards, and the entities of value at risk. It is important to note that the principal risk management concern is the management of the risks posed to people, and in particular, third parties not associated with the aviation operations taking place.

[e]Through the grant of a Certificate of Waiver or Authorization (COA); see https://go.usa.gov/x5hXW/

B. Operational Risks

Whilst the primary and secondary safety hazards are common to all UAS operations, the primary risks to be managed depend on the particular sUAS and the associated CONOPS.

For instance, in the UFT example (Section III.B.1), flights occur within the NAS, in particular within a bounded OR enclosing a sparsely populated area in California's central valley. The OR includes built-up structures, power lines, railroads, and vehicular traffic. Additionally, it is located in the vicinity of a handful of small private use and public airports. The air traffic in the area comprises general aviation aircraft, which may/may not be equipped with operating transponders, as well as low-altitude agricultural aviation, the flight paths of some of which are known to transect the OR. Thus, third party risk comprises airborne collision risk with non-cooperative, non-participating CPA, and ground collision risk leading to damage to structures or harm to people. For the QGC example (Section III.B.2), the risks to be managed are those relating to the primary hazard of a midair collision (MAC) and the secondary hazard of forest fires and/or damage to infrastructure as a result of a ground impact. Here, the primary MAC threat is an encounter with a small general aviation aircraft operating under VFR.

Amongst the two examples, despite the commonality both in the fundamental hazards and in some of the types of safety risk (e.g., airborne collision risk), the nature of risk is different. Indeed, given that operations are occurring in the vicinity of airfields and known VFR traffic routes in the UFT example, the likelihood of encountering air traffic, i.e., the (prior, unmitigated) probability that there is an airborne threat when the sUAS are also airborne, is likely to be greater than in the QGC example. Likewise, in the QGC example, the large geographical area of the mission, the increased sparsity of the ground population, will likely decrease the (prior, unmitigated) likelihood component of the risk of harm from a ground collision, in comparison with the UFT example, where the OR is substantially smaller. On the other hand, given the greater operating altitude and the larger weight of the sUAS in the QGC example, the impact from a ground collision is more likely to be lethal, therefore increasing the severity component of ground collision risk, compared to the UFT example.

C. Establishing a Risk Basis

In general, the components of ground collision risk involve exposure of the population to a ground impact, and the conditional probability that an impact leads to a fatality.[3,4] The former, in turn, depends on population density, the size of the crash area, the availability of shelter, and the degree to which shelter provides protection. The likelihood of an impact leading to a fatality is dependent on the kinetic energy transferred.

Similarly, the factors affecting airborne collision risk involve the exposure of CPAs to an airborne UA, and the conditional probability that an impact results in a hull loss (assumed to be fatal).[3] Exposure, in turn, is dependent on additional factors including air traffic density and the relative velocities of the aircraft in conflict. The former depends on the volume of the airspace within which the UA is operating, the relative sizes of the aircraft in conflict, and the number of aircraft in the airspace, whereas the latter is affected by factors such the encounter geometry. Each of these are further affected by characteristics of the prevailing air traffic management system, including airspace class, operating altitude, operating rules and procedures.

The factors affecting ground and airborne collision risk can be viewed as opportunities for risk management, which impact the additional safety measures required. For instance, in the QGC example, the nature and size of the mission area are likely to make it impractical to deploy and use ground-based primary surveillance radar, due to which multiple strategic and tactical safety measures may be preferable, in addition to other generic MAC risk controls.[9] More generally, the difference in risk between different CONOPS highlights the differences in the suitability of the risk controls (i.e., operational processes and technologies) available to manage them. Next, we describe how barrier models, in particular BTDs, provide an intuitive approach to safety risk management.

D. Bow Tie Modeling

Bow Tie Diagrams (BTDs), also known as *Bow Tie Models* (BTMs) or *Barrier Bow Tie Models* (BBTMs), represent a *barrier model* of safety, and provide a graphical means to visualize and assess the risk scenarios associated with a given hazard. Amongst the different approaches available for risk modeling in aviation, BTDs have seen extensive use, and they have been employed by national aviation regulators, air navigation service providers, and in defense aviation.[17,22,30–32] In the context of UAS, they have been used to enable non-segregated airspace operations[33,34] and, now, are also being recommended for the risk assessment of specific sUAS operations.[5] The authors have also successfully applied BTDs in practice, for sUAS risk modeling and in safety assurance of real operations.[8,9,12,35]

We do not describe the methodology or the steps to develop BTDs in this paper, and refer the reader to our previous work[9,12] as well as related literature.[30] In brief, the main components of a BTD (Fig. 2) are:

Hazard: A controlled activity, condition, or entity that reflects a normal or desirable aspect of the CONOPS, but can potentially be a source of harm if control is lost.

Top Event: An undesired system state, where there is a hazard release or a loss of control. We develop BTDs around a single top event associated with an identified hazard. For a MAC hazard, a possible top event is a failure to remain *well clear* of other aircraft.

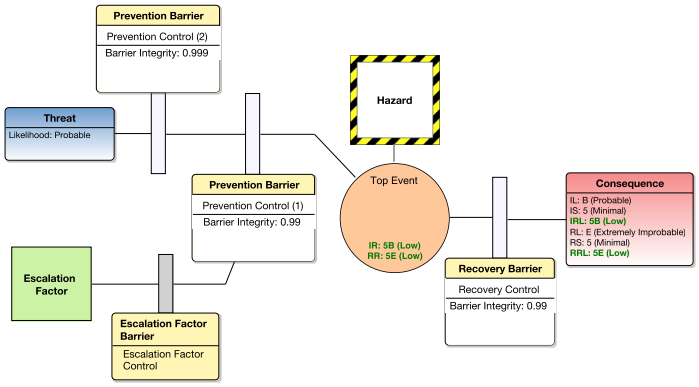


Figure 2. Components and structure of a generic Bow Tie Diagram.

Threat: A possible direct cause/source of the top event that generates it. Threats can include possible failure modes. Various general classifications of threats have been previously proposed for the sUAS primary hazards.[35]

Consequence: The potential dangerous outcome or loss state resulting from the top event that must be avoided, e.g., serious or fatal injury to a third party.

Control: Any process, device, practice, or other action that modifies risk.[36] Controls can be classified as *i*) *prevention*, or *preventative*, when they contribute to reducing the occurrence probability and/or the magnitude of severity of the top event, and *ii*) *recovery*, or *mitigative*, when they contribute to reducing the occurrence probability and/or the magnitude of severity of the consequence, given that the top event has occurred.

Barrier: A collection or *system* of controls that contributes to reducing the probability of occurrence and/or magnitude of severity of the consequence(s) associated with a particular event within a chain of events describing a risk scenario. Barriers can be ascribed a measure of *integrity*, that relates to the likelihood of a dangerous failure[f] of the barrier.

Escalation Factor: A weakness/vulnerability, threat, or operational condition that can compromise, defeat, or otherwise degrade control effectiveness. These can include environmental conditions, e.g., adverse weather.

Escalation Factor Barrier: A *second tier* or secondary system of controls used to manage, reduce or modify the impact an escalation factor has on another control.

Fig. 3 shows a fragment of a BTD for the UFT example, where the hazard (H1) consists of airborne UAs operating BVLOS within the OR. The top event of interest is a loss of safe separation between a UA and a CPA (E4), and its proximate causes are the identified threats—airborne excursion (E1), and airborne intrusion (E3). The consequence event to be avoided is a MAC (E7). This BTDs shows two *risk scenarios*, i.e., the distinct paths in the BTD each beginning at an initiating threat, and terminating at a particular consequence. Each risk scenario across the different BTDs applicable for a given CONOPS will need to be assessed and managed as part of the standard risk management process.

One or more barriers can be employed for each scenario to prevent a hazard/top event or otherwise mitigate its potential consequences. For example, here, *independent flight abort* is a preventative barrier applicable for the scenario beginning with an airborne intrusion, whereas a suite of *emergency procedures* would be the barrier invoked in the scenario beginning with an airborne excursion. Note that in Fig. 3, almost all of the barriers shown comprise exactly one control; although, more generally, a barrier can comprise multiple individual controls. We see this, for example, in the barrier labeled *piloting safety actions*, which consists of both a prevention and a recovery control (shown in Fig. 3, on either side of the top event E4).

It is, thus, intuitive to see how a BTD concisely conveys the planned treatment of risk scenarios associated with a hazard, through the use of controls/barriers and escalation factor barriers (not shown in Fig. 3). Similar BTDs can be constructed for other top events (if applicable), and for the remaining primary hazards of an sUAS operation (see Fig. 1), as appropriate. Although not in scope for this paper, it is worth noting that threat events can also be analyzed and managed in a similar way, by shifting the focus farther back in the event chain. For example, in Fig. 3, we can further analyze the airborne intrusion event (E3) by designating it as a top event in a new BTD, identifying its threats, and using barriers/controls to manage those threats.

[f]We adopt this term to distinguish it from the more familiar notion of *reliability*, which is concerned with *all* barrier breaches rather than those that are dangerous, and affect safety.

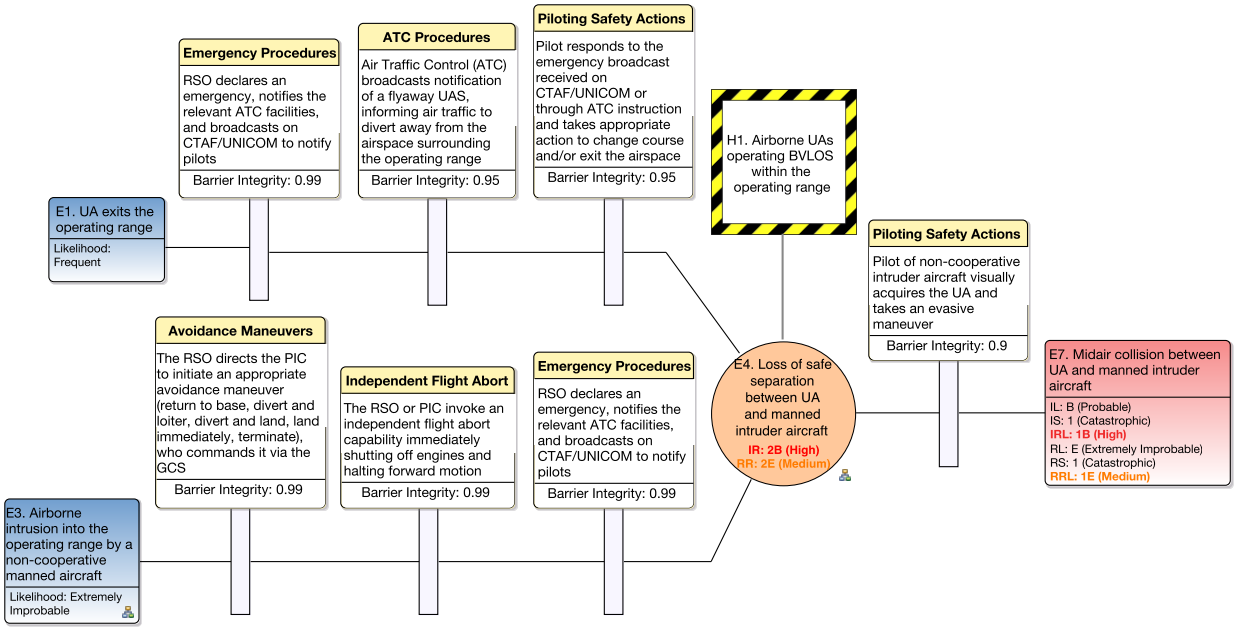


Figure 3. Fragment of one of the BTDs applicable to the CONOPS of the UFT example.

Risk assessment—more specifically, determining residual risk—involves using the threat probabilities (themselves determined from data or analysis), along with barrier effectiveness (or integrity), to qualify the likelihood of occurrence of the consequence(s). Care must be taken in this assessment to account for different risk scenarios that may share consequences.[37]

As we shall see subsequently in the paper, BTDs can be related to the various requirements of a particular safety decision making framework (see Section II.B), making it useful in the substantiation of claims against a safety objective. Additionally, BTDs have the added advantage of providing an overarching framework for other related techniques, including fault tree and event tree analyses, software assurance, and human factors analysis.

V. Safety Rationale and Assurance

A. Preliminaries

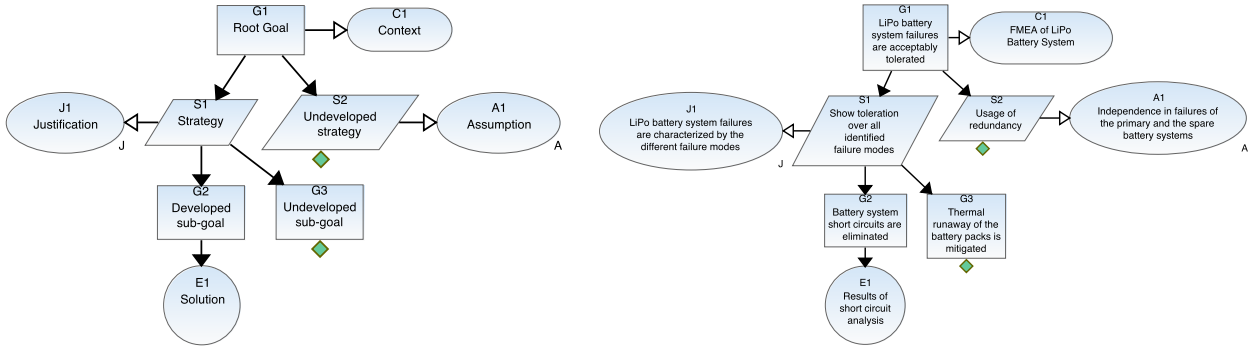
As mentioned earlier (Section II), the aim of a RISC in our context is to assure the regulator that sUAS operations can be safely conducted. In the RISC, this takes the form of specific high-level safety objectives—determined, in part, based on the applicable safety decision making framework (see Section II.B)—that are substantiated through structured argumentation. That is, we *i*) link operational safety to specific safety objectives that are, themselves, linked to the appropriate evidence, and *ii*) explicitly elaborate the rationale why the regulator should conclude based on the evidence provided that the safety objectives have been met.

One such safety objective concerns the reduction of risk to an acceptable level. To our knowledge, the literature provides limited guidance on how BTDs can be used to provide assurance of having met this safety objective. One of the challenges lies in the inherently qualitative nature of the risk assessment afforded by BTDs. Consequently, additional confidence must be engendered in the risk reduction purported to have been achieved by using the safety system (as modeled using BTDs). Whilst it is feasible to provide a simple and useful underlying model for quantitative analysis, the tradeoff is reduced accuracy that results in greater uncertainty in the risk assessment. Quantification and subsequent validation can, itself, present additional challenges.[37]

Next, (in Section V.B) we give a brief overview of structured argumentation, and how it captures (safety) rationale. Then (in Section V.C), we present a tiered framework for assurance, elaborating various assurance qualities/concerns. achievement of the safety objectives.

B. Structured Arguments

An *argument* is a connected series of propositions used in support of the truth of an overall proposition. We refer to the latter as a *claim*, whereas the former represents a chain of reasoning connecting the claim and the *evidence*. We do not describe the methodology or steps to create structured arguments in this paper and refer the readers to our previous work.[38]



(a) Core Goal Structuring Notation (GSN).

(b) Example GSN argument fragment.

Figure 4. Graphical presentation of a structured argument using GSN.

Structured arguments can be graphically depicted as a directed acyclic graph of different nodes and links, e.g., using the Goal Structuring Notation (GSN)[39] as shown in Fig. 4. The *core* GSN (Fig. 4a) comprises six node types—i.e., *goals*, *strategies*, *contexts*, *assumptions*, *justifications*, and *solutions*—and two link types that specify, respectively, *support* (→) or *contextual* (⇾) types of relationships between the nodes. The GSN standard also includes notational extensions for modularity,[39] though we will not cover those here.

In general, nodes refer to external items including *a*) artifacts such as hazard logs, requirements documents, design documents, various relevant models of the system, etc., *b*) the results of engineering activities, e.g., safety, system, and software analyses, various inspections, reviews, simulations, and verification activities including different kinds of system, subsystem, and component-level testing, formal verification, etc., and *c*) records from ongoing operations, as well as prior operations, if applicable.

Fig. 4b is an illustrative example of a GSN argument fragment (retaining the layout of Fig. 4a to aid understanding) to substantiate the main claim (goal node G1) of acceptably tolerating failures for a Lithium-polymer (LiPo) battery system, in the context (node C1) of its failure modes and effects analysis (FMEA). The structure below node G1 elaborates the rationale for accepting the main claim. Specifically, the argument uses two complementary strategies, i.e., S1: showing that all identified failure modes are tolerated, and S2: using redundancy. The latter relies on an assumption (node A1) of independence in failures of the redundant systems, but has not been further developed (as indicated by the '◇' node decoration). The justification (node J1) for the former is based in the assertion that the different failure modes characterize the overall failure behavior. One of those failure modes concerns short circuits, whose elimination (node G2) is shown using the results of a short circuit analysis (node E1). Another failure mode pertains to thermal runaway, whose mitigation (node G3) is yet to supported by evidence (again, indicated by the '◇' node decoration).

C. Tiered Assurance Framework

Table 1 presents a tiered assurance framework applicable for sUAS RISC, and identifies several safety objectives that are *core assurance concerns*: for instance, the overall system safety objective that is the focus of this paper relates to safety in operations, i.e., showing risk reduction to an acceptable level. Additional safety objectives that apply, as appropriate, include the broader aspects of system safety (e.g., showing that the implementation of the system will not compromise safety), showing compliance to regulations, and those arising from the safety decision making frameworks (e.g., showing that risks have been reduced to ALARP levels).

Besides the core concerns, safety objectives can arise from *additional assurance qualities*, though we do not address those objectives in this paper. Their scope includes qualities that *a*) could be used to show assurance in the safety/systems engineering and risk assessment process itself, i.e., assurance in the input data, people, and processes used, which can be considered as a *backing evidence* in a RISC,[17] and *b*) are required to ensure that a safety objective

Table 1. Tiered framework for assurance through a RISC.

Tier	Core Assurance Concerns and Scope			Additional Assurance Qualities
Safety Objectives	**System Safety** – Safe concept (safety designed-in) – Safety in design – Safety in implementation – Safe transition into service – **Safety in operations** – **TLOS / Acceptable level of risk** – Safe disposal	Due diligence Reduction of risk – ALARP – SFAIRP – ASARP	Compliance with Aviation Regulations	Processes; – Maturity, … Input data; People; – Competence, … Method and Tools; – Qualification, … Safety management system; Lifecycle
1	**Overall Assurance** All hazards / hazard risk statements, i.e., combination of hazardous situation, hazard release. **All relevant consequences** across all BTDs. (Fig. 5)		All applicable regulatory requirements	Coverage; Independence of threats; Effectiveness; ….
2	**Profile of Risks** For each hazard, all risk scenarios (consequences), e.g., midair collision, near midair collision, ground collision, … **Specific consequence,** e.g., **midair collision** (Fig. 6) All causal chains, threats, and dangerous interactions across all hazards.			Coverage (function, environment, interactions, scenarios, …); Independence; …
3	**Individual Risks** **Specific risk scenario,** i.e., causal chain of consequence, top event, threats, causes/precursors (Fig. 7) **Applicable system of barriers / safety measures** (Fig. 8)			Depth; **Independence** (Fig. 9); Proactiveness: Prevention vs. Recovery; …
4	**Barriers** Functional safety / **fitness for purpose** (Fig. 10) Delivery of required service			Depth; Independence; Common causes/modes, …
5	**Controls** Functional safety / fitness for purpose Delivery of required service			Reliability and effectiveness; Availability; Functional / safety integrity; Resilience; Fail safety; Data integrity; **Verifiability** (Fig. 11); …

is maintained across the operational lifecycle. The inspiration for the qualities listed in Table 1 was found through exploring general guidelines on aviation design, system safety, and quality assurance processes.

We now describe the various tiers at which assurance can be provided to bolster confidence in the RISC, elaborating on the different assurance qualities applicable at those tiers. We also describe the relationship of these qualities to the components and structure of BTDs. We will exemplify the discussion with fragments of GSN arguments, based on the UFT example (Section III.B.1) and its BTDs (of which Fig. 3 is an example fragment).

1. Overall Assurance

The interpretation of Table 1 is as follows: each of the five tiers successively refine the safety objectives related to the core assurance concerns in a top-down manner. For example, to show that the operational safety objective has been met, we provide assurance in the overall RISC, i.e., that all hazards have been effectively managed. A given sUAS CONOPS can induce multiple hazards, each of which can potentially have many top events. As such, a RISC can have many BTDs, each associated with a specific hazard/top-event pair, and each potentially having different consequences. Thus, the core assurance concern is equivalent to showing that the risk posed by the different consequences across all the relevant hazards and BTDs have been effectively managed.

Fig. 5 shows this rationale for the UFT example, where the main operational safety objective (goal G2) is linked to the risk posed by the identified consequences (goals G4: MAC, and G7: ground collision, respectively). One of the consequences, MAC, is exactly the consequence of the BTD in Fig. 3. The rationale captures the relevant assumptions, and contextual information, also showing some of the related safety objectives (e.g., goal G6, concerning implementation safety). To improve confidence, we must also provide assurance with respect to the additional assurance qualities corresponding to a tier. For Tier 1, for instance, one of the additional assurance qualities concerns *coverage* (e.g., of all consequences, all BTDs, and all hazards). An alternative argument structure (not shown here) could address each hazard, related top events, and their associated risk profiles, i.e., the BTD associated with top event. Here, the core assurance concern is showing effective management of the profile of risks for each BTD, i.e., Tier 2.

2. *Assurance in the Management of the Profile of Risks*

As indicated in the preceding discussion, the core assurance concern at this tier, i.e., showing effective management of the profile of risks, concerns the individual BTDs for each pair of identified hazard and top event. More specifically, we must provide assurance that all the risk scenarios of a given hazard have an acceptable residual risk level. That, in turn, requires showing that the individual risk scenarios have been effectively managed.

In the general case, different hazards or top events can lead to a common consequence. Moreover, since a consequence is the terminal event in a risk scenario, the core assurance concern is, equivalently, to show that, given a specific consequence, all the causal event chains leading to that consequence have been effectively managed.

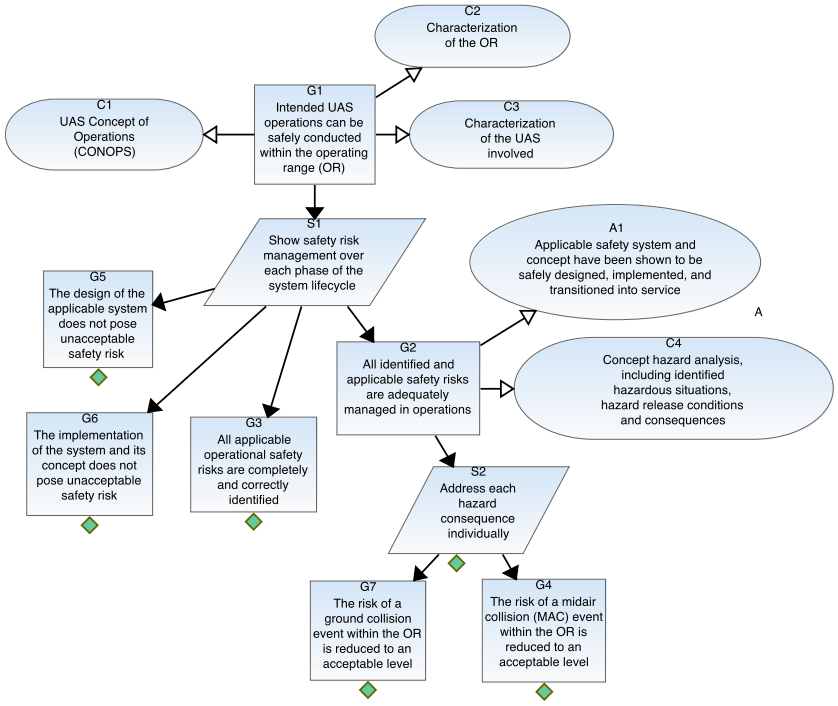


Figure 5. Tier 1, fragment of GSN assurance argument for the UFT example: linking operational safety to managing consequence risk.

Fig. 6 shows a fragment of the Tier 2 GSN assurance argument for the UFT example, reflecting this reasoning. Here, the root claim (goal G1)—that MAC consequence risk has been effectively reduced to an acceptable level—is both one of the core assurance concerns[g] of Tier 2, and a leaf claim of the argument in the previous tier (Fig. 5, goal G4), illustrating the link between the safety assurance rationale between the two tiers. We provide assurance by showing the reduction in likelihood of each corresponding, credible risk scenarios (Fig. 6, goals G4 and G5, respectively) that lead to a MAC consequence, i.e., the core assurance concern of Tier 3.

Additional assurance qualities at this tier (see Table 1) concern *independence* (e.g., in how risk scenarios are managed, amongst the scenarios themselves, as well as how the potential occurrence of multiple scenarios is managed), as well as *coverage* (that all applicable risk scenarios have been identified and treated).

3. *Assurance in the Management of Individual Risks*

The core assurance concerns at Tier 3 relate to the effective management of individual risk scenarios, i.e., each path or casual chain linking a specific threat to an individual consequence state within a BTD. At this level of abstraction, effectiveness of risk mitigation aggregates the effectiveness (equivalently, integrity) of the implemented barriers. To evaluate residual risk against the safety objectives, we (probabilistically) combine barrier integrities with the prior, unmitigated probability of occurrence of a threat.[37] Assurance in the management of the risks for a particular scenario is determined with respect to the integrated *system of barriers* (i.e., those in place between the corresponding threats and their consequences).

Fig. 7 shows a fragment of GSN assurance argument of the UFT example, from Tier 3. It continues the safety rationale from Tier 2 (Fig. 6), and it provides assurance of the reduction in likelihood of all credible risk scenarios leading to a MAC consequence. The assurance rationale uses two complementary lines of reasoning: first, probabilistic modeling shows risk reduction (as discussed above, through the combination of barrier integrities and threat probability), given by the argument leg G5 → S1 → G6 → E1. Next, mitigation of each applicable risk scenario is shown (see the argument fragment below strategy S2) by elaborating the safety barriers used in the applicable BTD (Fig. 3).

[g]The other core assurance concerns at this tier pertain to the remaining identified consequences/risk scenarios.

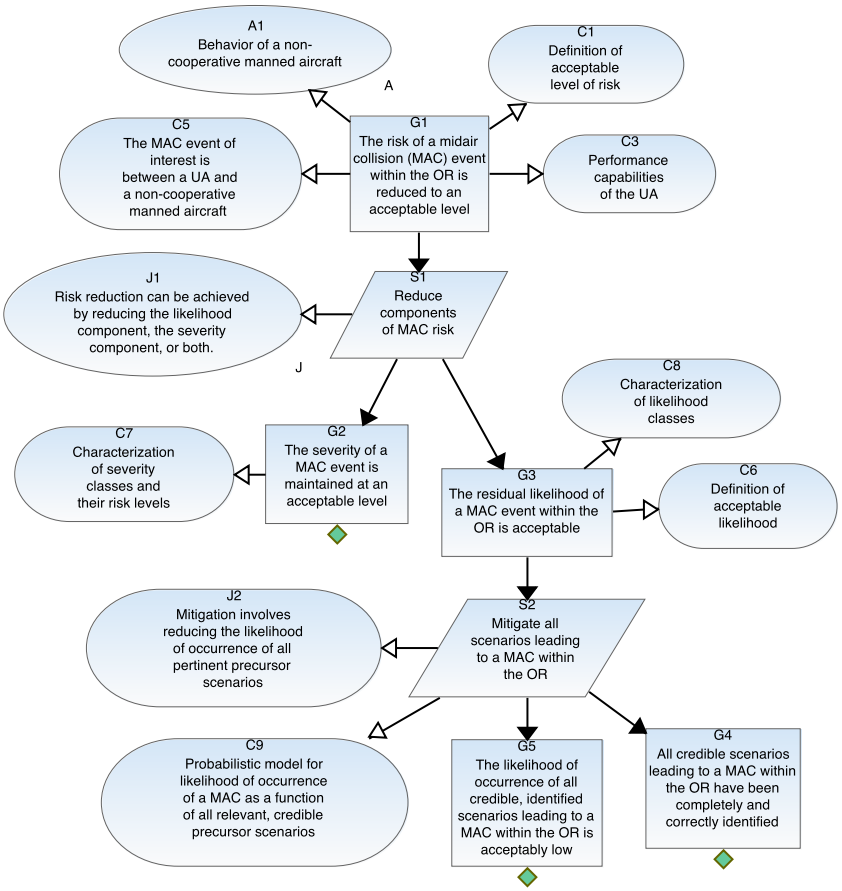


Figure 6. Tier 2, fragment of GSN assurance argument for the UFT example: continuing the Tier 1 rationale (Fig. 5), linking effective management of MAC risk to mitigation of all applicable risk scenarios.

That, in turn, requires showing that each initiating threat (of the scenarios under consideration) is, itself, effectively managed. For example, one such threat (Fig. 7, goal G14) concerns airborne intruders.

Fig. 8 continues the safety rationale of Fig. 7, and shows how the threat posed by an airborne intruder is managed by invoking a system of (prevention) barriers (see Fig. 3). The key observation here is that in this assurance argument, the specific contribution that each barrier makes to effectively manage a MAC risk scenario, is made explicit by the safety mechanisms provided to manage the threat that initiates the scenario. For example, a surveillance barrier (solution E1) provides sufficiently early warning (goal G15), while the avoidance maneuvers invoked (solution E2) provide one of two redundant means for separating the UAs from an intruder (goal G20).

At this tier, additional assurance properties that can be used to infer that the core assurance concerns have been addressed include *depth*, *independence*, and *proactiveness*. The former two together provide confidence that the system of barriers used to treat a risk scenario are sufficiently reliable and effective.

Depth: This property characterizes the extent or degree to which protections, i.e., barriers, are layered. Intuitively, the greater the depth, the greater the degree of risk reduction.

Independence: Barriers are (stochastically) independent if the event that one is breached does not affect the probability that another is also breached. Since common mode/cause failure effects are an indication of dependencies between barriers, this property characterizes the degree to which such failure causes or modes are absent. Fig. 9 gives a fragment of a structured GSN argument providing assurance that two specific barriers, namely those used for surveillance and avoidance, are independent when used in the risk scenario concerning a MAC consequence. The argument shows four complementary legs of reasoning, one of which—as mentioned earlier—appeals to the management of common causes/failure modes, while the others invoke assurance strategies including implementation and data diversity. Note that this fragment is, in fact, part of a wider argument (not shown) for assurance of independence in the system of barriers used to mitigate specific risk scenarios.

Proactiveness: This describes the extent to which risk management preference for using prevention (or preventative) barriers is met. Compared to mitigative / recovery barriers, preventative barriers reflect a more proactive treatment of risk. Prevention barriers aim to interrupt the risk scenario before the realization of the top event. Evidence of this property can be observed within the BTD as the number of preventative barriers implemented.

4. Assurance in Barriers

The core assurance concerns in Tier 4 pertain to the specific barriers deployed and their effectiveness in mitigating risk scenarios, in particular their fitness for purpose, and their functional safety. Barrier effectiveness (or integrity) pertains to its efficacy at reducing risk. As mentioned earlier, one or more controls comprise a barrier, and barrier effectiveness is an emergent property arising from the integrated behavior of its constituent controls.

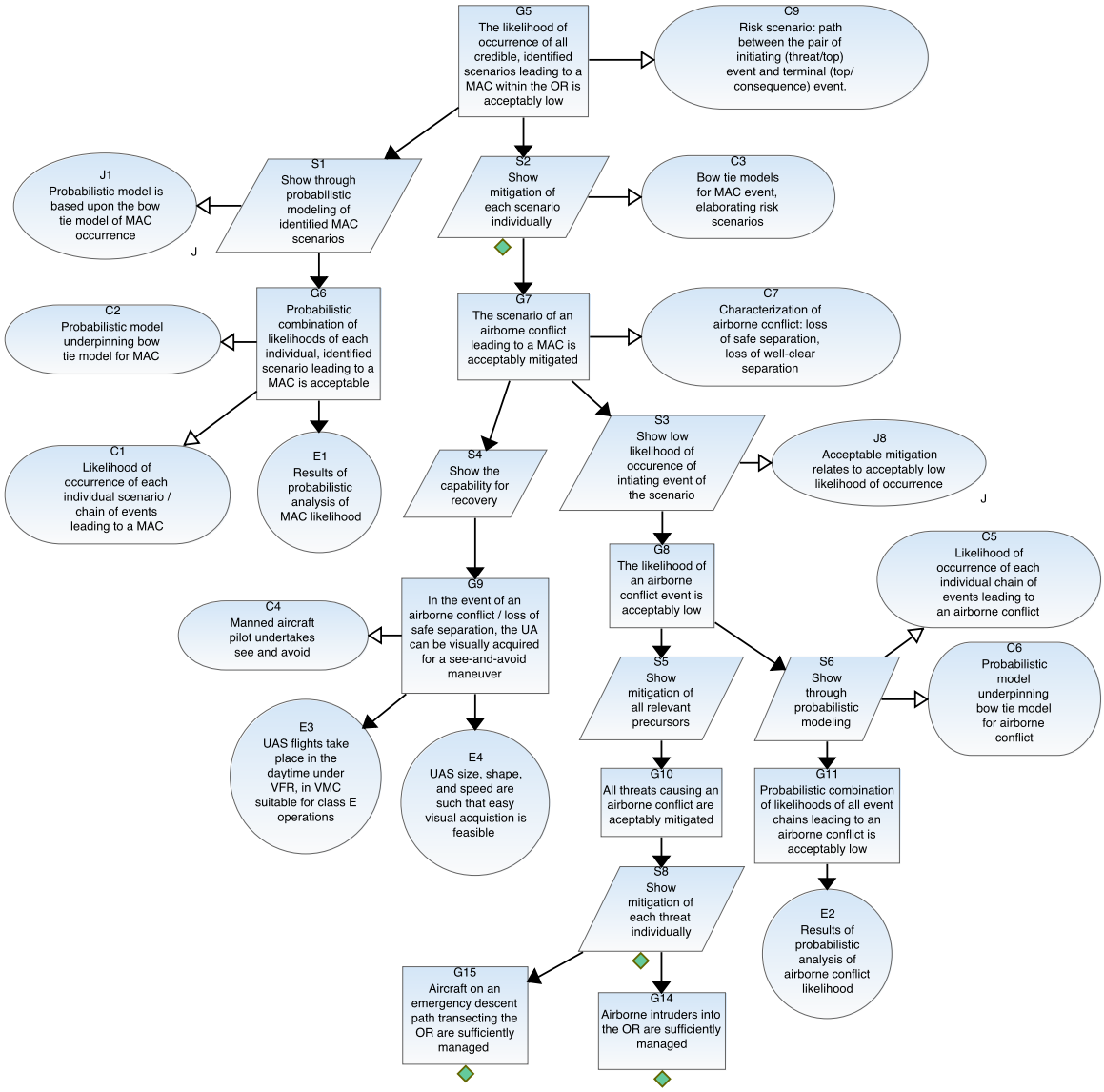


Figure 7. Tier 3, fragment of GSN assurance argument for the UFT example: assurance of risk reduction of individual risk scenarios leading to a MAC, continuing the safety substantiation rationale from Fig. 6.

Fig. 10 shows a fragment of a GSN assurance argument from Tier 4, that addresses the fitness for purpose of the surveillance barrier. The main assurance claim (goal G2), pertains to the surveillance infrastructure delivering the required service (i.e., detecting and tracking both cooperative and non-cooperative air traffic). Note that this is also a supporting claim in the argument at Tier 3 (Fig. 8, goal G22), providing assurance that the threat posed by an airborne intruder in a MAC risk scenario is effectively managed. As shown in Fig. 10, the main approach to demonstrate this assurance concern is a refinement of the safety requirement and showing iteratively that each applicable lower-level requirement is supported by evidence.

The additional assurance properties that apply at this tier (Table 1) include, as earlier, *depth* and *independence*. These properties are analogous to those applied to the system of barriers that mitigate risk scenarios, although for individual barriers, they refer to the depth / independence of the constituent controls.

Depth and independence (in controls) are, thus, strategies used for improving barrier reliability (thereby, its safety integrity). Barrier depth relates to the design principle of defense-in-depth, which can be measured by the number of unique controls comprising a barrier. However, the number of controls on its own is not sufficient for demonstrating reliability in a barrier. There can be numerous common mode failures between controls which, when realized, can degrade or defeat a barrier. Consequently a barrier should also exhibit the property of independence.

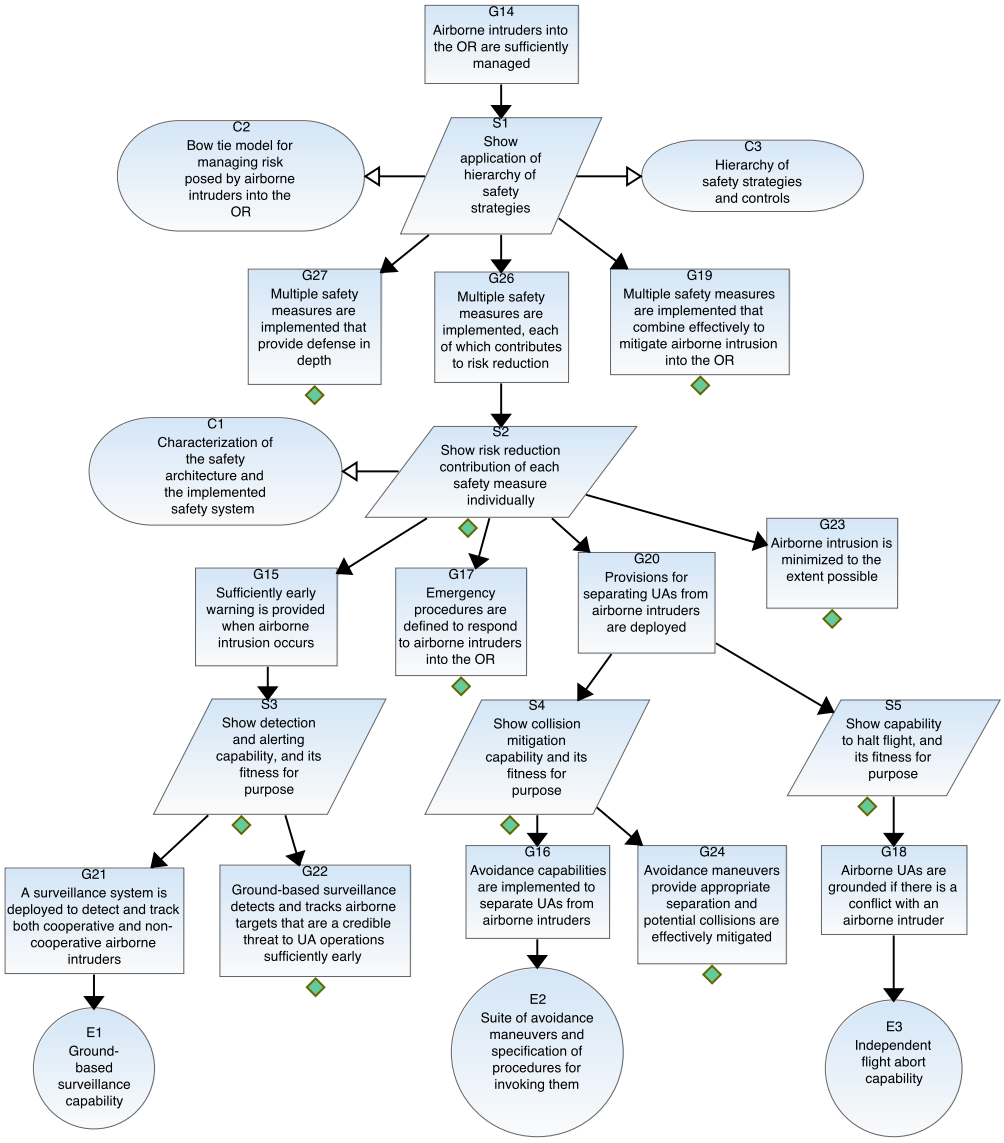


Figure 8. Tier 3, fragment of GSN assurance argument for the UFT example, continuing the rationale of Fig. 7: assurance of risk reduction of a particular risk scenario by managing a specific threat, i.e., airborne intruders.

The normal operation of one control could degrade that of another. Similarly, the failure of a control can cause the degradation of, or failure of, another. Common mode failure analysis techniques[40] can be used to identify such conditions and in turn, used as evidence of independence. Independence in barriers can also be observed within the BTDs by the absence of common controls and common escalation factors between controls belonging to a particular barrier.

5. Assurance in Controls

Assurance provided at Tier 5 represents the lowest level of abstraction, and addresses individual controls within a barrier. Controls can have varying levels of effectiveness. Effectiveness in this context describes the efficacy of a control at reducing risk; encompassing the magnitude of the reduction of risk it achieves and the manner in which it is achieved. Some controls are deemed more effective or preferred over others based on the manner in which they achieve the risk reduction. Risk management guidelines codify this as the *hierarchy of controls*. Assurance at this level of abstraction relates to the degree of confidence that a control is effective, i.e., it exhibits a high degree of (safety) integrity. Aviation safety literature define numerous general qualities that can be adapted and applied to the problem

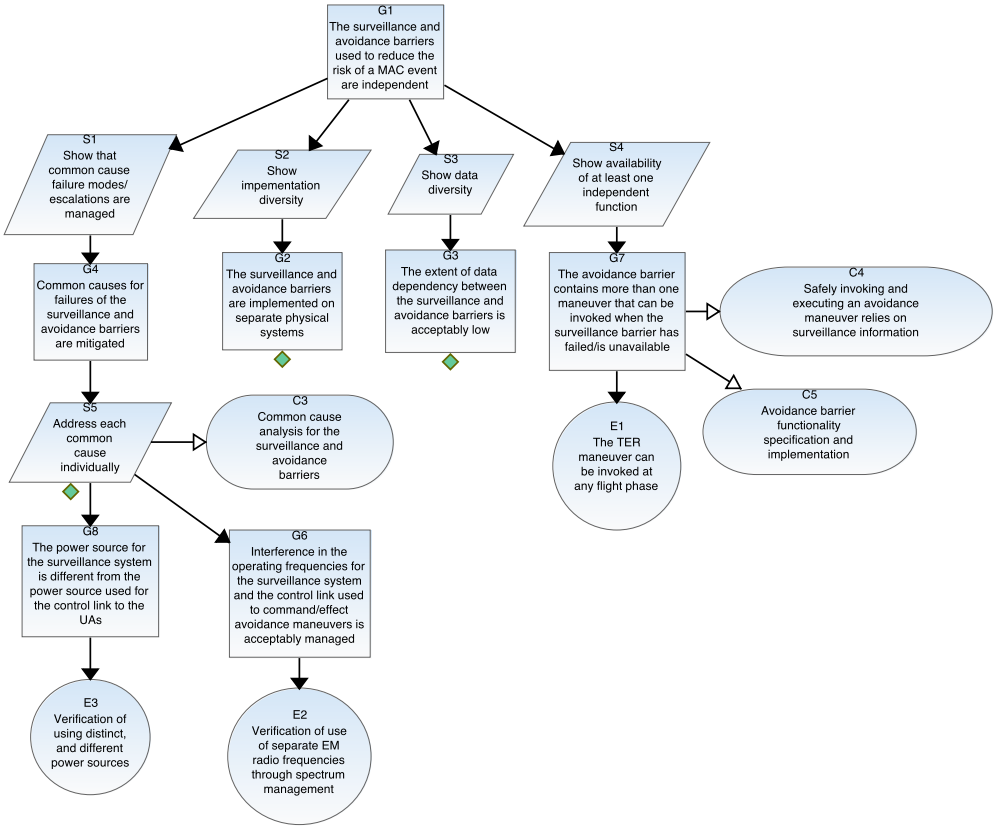


Figure 9. Tier 3, fragment of GSN assurance argument for independence of the surveillance and avoidance barriers.

of providing assurance in individual risk controls. Identified qualities include, but are not limited to:

Reliability: The probability that a control will perform its required function under specified conditions, without failure, for a specified period of time. A control must be both effective and reliable. Evidence of the reliability of a control can be provided through analysis (e.g., fault tree analysis), data from testing, standards on design, and testing, etc. With respect to the components of a BTD, reliability can also be evidenced through the presence of escalation barriers for all escalation factors associated with a given control.

Availability: The probability that a control is ready for use or in a functioning state at a given point in time. A control should also exhibit the property of availability. Availability takes into consideration that not all controls may not be effective over the entire duration or scope of a mission. For example, some controls may not be available at night or during those phases of a mission where the UA is BVLOS of its control station. In such cases the control is not able to perform its function but has not failed. An operator must also know when a control has failed or when not to use a control.

Functional Integrity: The probability that a control completes its intended function with no undetected error, or if there is an error, the probability that the error will be detected and a usable integrity flag generated within a specified maximum time[h]. A control exhibiting high functional integrity provides the operator assurance that the control will only be used when it is free from fault or error, or otherwise when its not appropriate to use the control. Features of a control that can be used as evidence of integrity include built-in-test, inspection, warning, and alerting functions.

Resilience: The probability a control can accommodate a failure within acceptable degradation parameters and to recover within an acceptable time. Resilience describes the ability of a system to continue functioning despite the occurrence of failures.

Fail Safety: The probability that, given a failure, a control can attain an operational state that mitigates the potential consequences of its failure. The property of fail safety is concerned with managing the impact of a loss of a control (e.g., its impact on other controls or creation of new risks).

[h]This definition is based on that provided by Sabatini et al.[41]

Figure 10. Tier 4, fragment of a GSN argument for the UFT example: addressing the fitness for purpose of the surveillance barrier, in its contribution to risk reduction.

Verifiability: The ability of a control to be checked or audited (by a person, tool, or other means) to determine whether it is operational and/or correctly implemented. This is closely related to the concept of enforceability, which is described as "the extent to which compliance with new rules, regulations or operating procedures can be monitored".[42] A verifiable control is one that is accessible and observable during its operation. Fig. 11 shows a fragment of a GSN assurance argument the provides the rationale to justify how the performance and effectiveness of a specific control pertaining to the surveillance barrier, i.e., a ground-based primary surveillance radar, is verifiable.

Additionally, the properties of resilience and fail safety relate to the behavior of a control given a failure or disruption. In addition to data from testing, evidence of the resilience and fail safety properties of a control can be provided through the adoption of sound engineering design principles (i.e., fault tolerant design) and documented system safety analysis.

These qualities are not intended to be comprehensive but are provided as examples of typical generic system properties that can be used to infer assurance in a control performing its intended risk mitigating function. There are also other qualities that can be important in the selection of controls but not directly related to assurance in the control. These include cost, practicality, acceptability of the controls, and introduced risk, where the acceptability of a control describes the extent to which the control is consistent with stakeholder paradigms[42] and introduced risk describes the magnitude of additional risk introduced to a system through the implementation of a control.

VI. Concluding Remarks

We have presented an approach to safety assurance of sUAS operations using *Risk Informed Safety Cases* (RISCs), which use BTDs to provide a risk basis, along with structured arguments for safety substantiating rationale.

BTDs are particularly useful during risk evaluation and treatment as they focus analysis and decision making on the mechanisms for controlling risk. Moreover, they can help to establish the relationships between the implemented risk controls, the mechanisms for assurance in the provision of the controls, and the consequences associated with the loss (or breach) of the controls across multiple possible accident scenarios.[35]

Structured arguments, given here using GSN, enable the explicit tracing of safety and airworthiness considerations, from concept, to requirements, to evidence of risk mitigation and control. They are useful as a centralized organizing component of the diverse assurance information aggregated in a RISC. Moreover, they facilitate conveying qualitative rationale for why various safety objectives, such as claims about effective safety risk management, risk reduction to acceptable levels, etc., are likely to have been achieved. Thus, they provide a basis with which regulators, and other stakeholders of sUAS operations can be justifiably confident of operational safety.

We have outlined various *core assurance concerns* and *additional assurance qualities*, where we believe arguments can provide insight and justification. These concerns can be organized naturally in a tiered framework, thereby providing a roadmap for structuring the safety case, whilst also serving as a useful catalogue of assurance qualities.

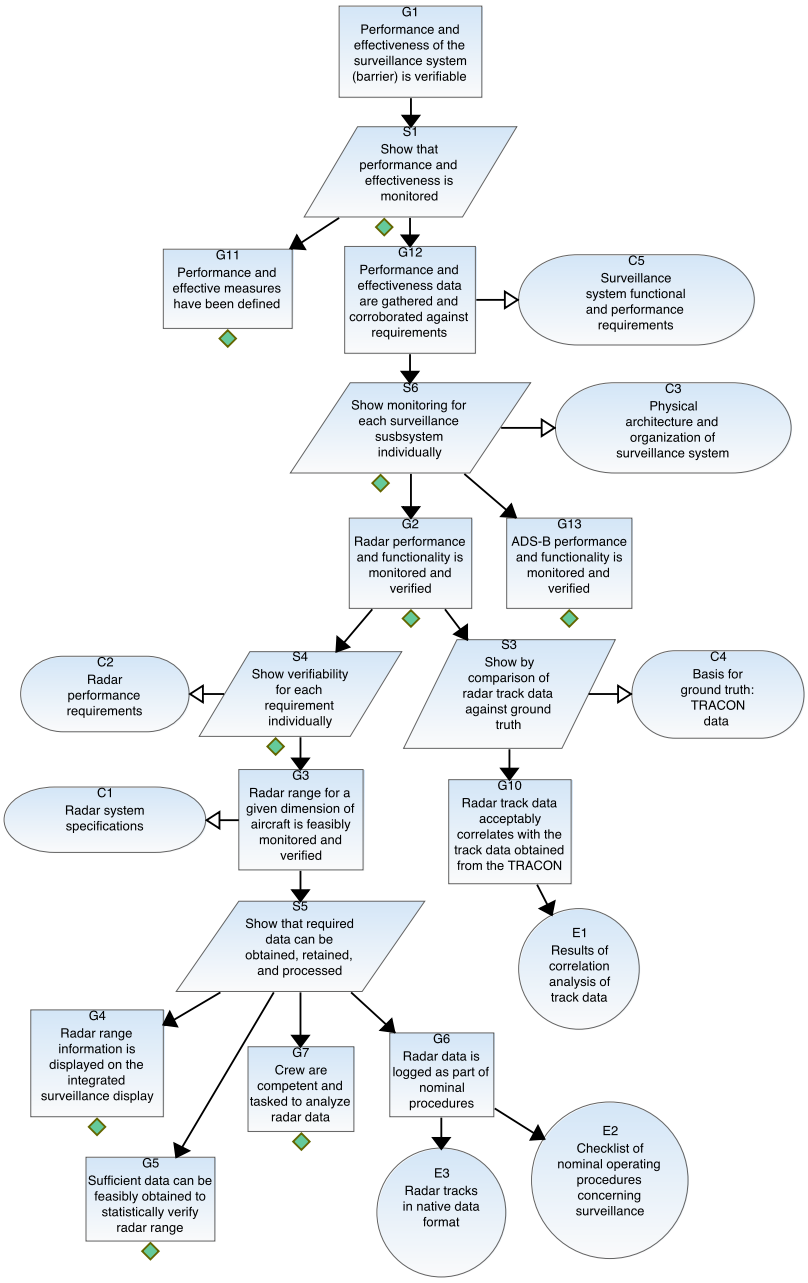


Figure 11. Tier 5, fragment of GSN assurance argument for verifiability of a specific control within the surveillance barrier, i.e., a radar system.

The identified assurance qualities do not necessarily need to be quantitatively assessed. They can serve as avenues of qualitative inquiry for a regulator (or operator). We believe that the qualities will help to bridge the gap between the higher level safety case, and the lower level requirements on the design, implementation, operation, and sustainment of controls.

We have described two example sUAS CONOPS that exhibit commonalities and differences both in the hazards posed and the associated safety risks to be managed, and we have illustrated our approach focusing on one of those examples. Though we have concentrated here on a midair collision (MAC) consequence, the same approach is applicable to the assessment of risk posed by other hazards, e.g., ground collision risk. In the context of assurance, whereas

the arguments from the higher-level tiers are expected to be largely common, and applicable to either CONOPS, the lower tiers will be specific to particular controls and barriers. In general, the use of BTDs with argumentation will be applicable to any system where a layered protection approach is used.

We have implemented facets of our approach—specifically, BTDs and assurance arguments—in the AdvoCATE toolset, which provides support for both GSN-based argumentation and BTD, and navigation between the two[37,43] and was used to create the diagrams in this paper.

In the context of the UFT example (which represents ongoing sUAS operations being conducted by NASA as part of the UTM effort) we note that a RISC (created by the second and third author) used BTDs for the key risk assessments. Additionally structured arguments provided the safety rationale for the critical barriers, where the regulator required additional assurance of fitness for purpose. Subsequently, this safety case successfully underwent regulatory evaluation, and the associated sUAS operations were granted operational approval to conduct BVLOS flight operations. Although, this safety case did not employ the tiered approach to assurance, which evolved from our respective experiences in developing large safety cases, we plan to use this framework as a guideline to structure subsequent safety assurance efforts.

Although a single overarching argument that integrates all of the assurance concerns could be given for the overall safety system, in practice, we have found it useful to focus on those components and aspects of the safety system where assurance is particularly required: for example, where a single barrier is used to manage a risk scenario, or where the independence of barriers (a key assumption in computing residual risk levels using BTDs) requires justification.

Moreover, whilst the tiered assurance framework outlined in this paper has identified various assurance facets, it has not yet elaborated when, or where it is appropriate to use arguments for conveying rationale. For instance, structured arguments would be useful in providing assurance in those situations where quantification is difficult or infeasible. On the other hand, when sufficient data is available and quantification is feasible—e.g., in the assurance of control or barrier reliability, availability, safety integrity, etc.—argumentation may be less appropriate. Likewise, arguments may be suitable for those barriers considered *critical*. Thus, we plan to investigate how the assurance framework and the underlying process can, itself, be made commensurate with the risk posed.

Acknowledgments

This work has been supported, in part, by the Safe Autonomous Systems Operations (SASO) project, under the Airspace Operations and Safety Program (AOSP) of the NASA Aeronautics Research Mission Directorate. We also acknowledge Mr. Brendan Williams, Boeing Research & Technology – Australia, and Ms. Kelly Hayhurst, NASA Langley Research Center, for their contributions during the initial development of some of the concepts presented in this paper.

References

[1]Subcommittee F38.01, ASTM International, "Standard Specification for Design and Construction of a Small Unmanned Aircraft System (sUAS)," ASTM Standard F2910-14, 2014.

[2]Special Committee (SC) 228, RTCA Inc., "Command and Control (C2) Data Link Minimum Operational Performance Standards (MOPS) (Terrestrial)," RTCA DO-362, Sep. 2016.

[3]Weibel, R. E. and Hansman, R. J., "Safety Considerations for Operation of Different Classes of UAVs in the NAS," *AIAA 3rd "Unmanned Unlimited" Technical Conference, Workshop and Exhibit, Infotech@Aerospace Conferences*, No. AIAA 2004-6421, Sep. 2004.

[4]Patterson, B., Lester, T., and Breunig, J., "Proposed sUAS Safety Performance Requirements for Operations over People," Presented at the 2016 MIT Lincoln Laboratory Air Traffic Control Workshop, Dec. 2016.

[5]Joint Authorities for Rulemaking of Unmanned Systems, "JARUS guidelines on Specific Operations Risk Assessment (SORA) (External Consultation Draft)," JAR-DEL-WG6-D.03, Aug. 2016.

[6]Subcommittee F38.02, ASTM International, "Standard Practice for Operational Risk Assessment of Small Unmanned Aircraft Systems," ASTM Standard F3178-16, 2016.

[7]Kunzi, F., "Framework for Risk-based Derivation of Performance and Interoperability Requirements for UTM Avionics," *Proceedings of the 35th IEEE/AIAA Digital Avionics Systems Conference (DASC 2016)*, Sacramento, CA, Sept. 2016, pp. 1–10.

[8]Williams, B., Clothier, R., Fulton, N., Lin, X., Johnson, S., and Cox, K., "Building the Safety Case for UAS Operations in Support of Natural Disaster Response," *14th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, No. AIAA 2014-2286, Jun. 2014.

[9]Clothier, R. A., Williams, B. P., and Fulton, N. L., "Structuring the Safety Case for Unmanned Aircraft System Operations in Non-segregated Airspace," *Safety Science*, Vol. 79, 2015, pp. 213 – 228.

[10]Denney, E. and Pai, G., "Argument-based Airworthiness Assurance of Small UAS," *Proceedings of the 34th IEEE/AIAA Digital Avionics Systems Conference (DASC 2015)*, Prague, Czech Republic, Sep. 2015, pp. 5E4–1–5E4–17.

[11]Denney, E. and Pai, G., "Safety Considerations for UAS Ground-based Detect and Avoid," *Proceedings of the 35th IEEE/AIAA Digital Avionics Systems Conference (DASC 2016)*, Sacramento, CA, 2016, pp. 1–10.

[12]Denney, E. and Pai, G., "Architecting a Safety Case for UAS Flight Operations," *34th International System Safety Conference (ISSC 2016)*, Orlando, FL, Aug. 2016.

[13]Dezfuli, H., Benjamin, A., Everett, C., Smith, C., Stamatelatos, M., and Youngblood, R., *NASA/SP-2010-580, NASA System Safety Handbook. Volume 1, System Safety Framework and Concepts for Implementation*, NASA, Nov. 2011.

[14]Dezfuli, H., Everett, C., and Groen, F., "The Evolution of System Safety at NASA," *International System Safety Training Symposium (ISSTS)*, St. Louis, MO, 2014.

[15]US Dept. of Transportation, Federal Aviation Administration (FAA), "Flight Standards Information Management System, Volume 16, Unmanned Aircraft Systems," Order 8900.1, [online] http://fsims.faa.gov/ [accessed May 2017], Jun. 2014.

[16]Eurocontrol, "Safety Case Development Manual," DAP/SSH/091 Ed. 2.1 [online], https://goo.gl/7FXYHp [accessed 15 May 2017], Oct. 2006.

[17]Safety Regulation Group, UK Civil Aviation Authority (CAA), "CAP 760 Guidance on the Conduct of Hazard Identification, Risk Assessment and the Production of Safety Cases for Aerodrome Operators and Air Traffic Service Providers," Dec. 2010.

[18]UK Civil Aviation Authority (CAA), "Small Unmanned Aircraft: Congested Areas Operating Safety Case (CAOSC)," Information Notice IN-2014/184, Nov. 2014.

[19]UK Ministry of Defence, "Safety Management Requirements for Defence Systems," Defence Standard 00-56 Part 1 Issue 7, Feb. 2017.

[20]Groen, F. J., Evans, J. W., and Hall, A. J., "A Vision for Spaceflight Reliability: NASA's Objectives based Strategy," *2015 Annual Reliability and Maintainability Symposium (RAMS)*, Palm Harbor, FL, Jan. 2015, pp. 1–6.

[21]Civil Aviation Safety Authority (CASA), "Common Risk Management Framework for Airspace and Air Traffic Management," [online], https://infrastructure.gov.au/aviation/airspace_reform/crmf.aspx [retrieved 15 May 2017], Oct. 2013.

[22]Defence Aviation Safety Authority (DASA), Australian Department of Defence, "Defence Aviation Safety Regulation (DASR), DASR Safety Management Systems (SMS), GM SMS.A.25 (b) (2) Safety Risk Management (AUS) - General," AAP 8000.011, Jan. 2017.

[23]Ale, B. J. M., "Tolerable or Acceptable: A Comparison of Risk Regulation in the United Kingdom and in the Netherlands," *Risk Analysis*, Vol. 25, No. 2, Apr. 2005, pp. 231–241.

[24]UAS Task Force Airspace Integration Integrated Product Team, "Department of Defense Unmanned Aircraft System Airspace Integration Plan, Version 2.0," OSD Report RefID: 1-7ABA52E [online], https://go.usa.gov/xNatN [accessed 15 May 2017], Mar. 2011.

[25]Eurocontrol, "RPAS ATM CONOPS," ATM.STR.CONOPS-RPAS.V(E) v4.0, Feb. 2017.

[26]Prevot, T., Rios, J., Kopardekar, P., Robinson III, J., Johnson, M., and Jung, J., "UAS Traffic Management (UTM) Concept of Operations to Safely Enable Low Altitude Flight Operations," *Proceedings of 16th AIAA Aviation Technology, Integration, and Operations Conference*, No. AIAA-2016-3292, Jun. 2016.

[27]US Dept. of Transportation, Federal Aviation Administration, AJV-115, Emerging Technologies Team, "Unmanned Aircraft Systems (UAS)," Air Traffic Organization Policy Order JO 7200.23, Aug. 2016.

[28]Insitu Pacific, "ScanEagle Unmanned Technology Benefits Commercial Sector in Natural Gas BVLOS Operations," [online press release], https://goo.gl/djNdDV [retrieved 15 May 2017], May 2016.

[29]Clothier, R. A. and Walker, R. A., *The Safety Risk Management of Unmanned Aircraft Systems*, chap. 92, Handbook of Unmanned Aerial Vehicles, Springer Netherlands, Dordrecht, Netherlands, 1st ed., 2015, pp. 2229–2275.

[30]UK Civil Aviation Authority (CAA), "Bowtie Risk Assessment Models," [online], http://www.caa.co.uk/Safety-Initiatives-and-Resources/Working-with-industry/Bowtie/ [retrieved 15 May 2017], 2015.

[31]FAA Air Traffic Organization, Safety and Technical Training Service Unit, "Transforming Risk Management: Understanding the Challenges of Safety Risk Measurement," [online], https://go.usa.gov/xXxea [accessed 15 May 2017], Dec. 2016.

[32]Acfield, A. P. and Weaver, R. A., "Integrating Safety Management Through the Bowtie Concept A Move Away from the Safety Case Focus," *Proceedings of the Australian System Safety Conference (ASSC 2012)*, edited by T. Cant, Vol. 145, Australian Computer Society, CRPIT, Brisbane, Australia, 2012, pp. 3–12.

[33]Eurocontrol, "Unmanned Aircraft Systems - ATM Collision Avoidance Requirements," CND/CoE/CNS/09-156, May 2010.

[34]Eurocontrol, "Air Traffic Management Guidelines for Global Hawk in European Airspace," EUROCONTROL Specification and Guidelines, Dec. 2010.

[35]Clothier, R., Williams, B., and Washington, A., "Development of a Template Safety Case for Unmanned Aircraft Operations Over Populous Areas," SAE Technical Paper [online], doi: 10.4271/2015-01-2469, Sep. 2015.

[36]International Organization for Standardization (ISO), Technical Committee 262, Risk Management, "Risk Management - Principles and Guidelines," ISO Standard 31000:2009, Nov. 2009.

[37]Denney, E., Pai, G., and Whiteside, I., "Modeling the Safety Architecture of UAS Flight Operations," *Proceedings of the 36th International Conference on Computer Safety, Reliability, and Security (SAFECOMP 2017)*, edited by S. Tonetta, E. Schoitsch, and F. Bitsch, Lecture Notes in Computer Science (LNCS), Springer, Sep. 2017 (to appear).

[38]Denney, E. and Pai, G., "A Methodology for the Development of Assurance Arguments for Unmanned Aircraft Systems," *Proceedings of the 33rd International System Safety Conference (ISSC 2015)*, San Diego, CA, Aug. 2015.

[39]Goal Structuring Notation Working Group, "GSN Community Standard Version 1," [online], http://www.goalstructuringnotation.info/ [retrieved 15 May 2017], Nov. 2011.

[40]SAE International, "Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment," ARP4761, 1996.

[41]Sabatini, R., Moore, T., and Hill, C., "A New Avionics-Based GNSS Integrity Augmentation System: Part 1 - Fundamentals," *The Journal of Navigation*, Vol. 66, No. 3, 2013, pp. 363–383.

[42]International Civil Aviation Organization (ICAO), "Safety Management Manual (SMM)," DOC 9859, AN/474, 2013.

[43]Denney, E. and Pai, G., "Tool Support for Assurance Case Development," *Automated Software Engineering*, 2017, To appear.