Risk-management of UAS Robust Autonomy for Integration into Civil Aviation Safety Frameworks

Tristan Perez[1] **Reece A. Clothier**[2] **Brendan Williams**[3]

[1] School of Engineering
The University of Newcastle,
Callaghan, New South Wales 2308, Australia
Email: tristan.perez@newcastle.edu.au

[2] School of Aerospace, Mechanical, and Manufacturing Engineering
RMIT University,
PO Box 71, Bundoora, Victoria 3083, Australia
Email: reece.clothier@rmit.edu.au

[3] Boeing Research & Technology Australia
GPO Box 767, Brisbane, Queensland 4001, Australia
Email: brendan.p.williams@boeing.com

Abstract

This paper discusses a model of the civil aviation regulation framework and shows how the current assessment of reliability and risk for piloted aircraft has a limited applicability for Unmanned Aircraft Systems (UAS) as technology moves towards higher levels of autonomous decision making. Then, a new framework for risk management of robust autonomy is proposed, which arises from combining quantified measures of risk with normative decision making. The term *Robust Autonomy* describes the ability of an autonomous system to either continue or abort its operation whilst not breaching a minimum level of acceptable safety in the presence of anomalous conditions. The term combines reliability, safety, and robustness. The decision making associated with risk management requires quantifying probabilities associated with the measures of risk and also consequences of outcomes related to the behaviour of autonomy. The probabilities are computed from an assessment under both nominal and anomalous scenarios described by faults, which can be associated with the aircraft's actuators, sensors, communication link, changes in dynamics, and the presence of other aircraft in the operational space. The consequences of outcomes are characterised by a loss function quantifies the desirability of the outcomes.

Keywords: Robust Autonomy, Unmanned Aircraft Systems, UAS, Regulation, Certification, Bayesian Reliability

1 Introduction

Unmanned Aircraft Systems (UAS) are a rapidly growing sector of the civil aviation industry. A key challenge facing this emerging sector is the lack of a regulatory framework, either prescriptive, performance or goal based, that can provide assurance in the safety of UAS operations. This framework must account for the varying Levels of Autonomy (LoA) in UAS.

There is an increasing demand for higher LoA in UAS. This demand largely stems from the need to lower the operational cost of UAS by reducing the number of people required to operate the system and to reduce the need for continuous communication links. The role of the human Remote Pilot (RP) in the operation of the UAS depends on the LoA of the system. The higher the LoA, the more the UAS subsumes the role (i.e., functions) of the human RP (Clothier & Walker 2013). Thus, as the LoA increases so does the complexity of the UAS and the higher the degree of safety assurance that must be demonstrated by the components of the UAS. This creates a new and challenging paradigm for National Airworthiness Authorities (NAAs) responsible for managing the safety of the civil UAS industry.

One of the first steps in the safety management process is the establishment of stakeholder goals in relation to the safety performance of the system. A risk assessment is then undertaken to estimate the safety performance of the system. The risk estimates are then compared against the goals to determine which of the identified hazardous scenarios require mitigation and control. In a civil aviation safety management context, the mechanisms put in place to reduce (mitigate and control) the identified risks establish the framework of regulations, standards and procedures relating to

1. design, manufacture, maintenance, and operation of the aircraft (i.e., UAS),
2. training and licensing of personnel, and the
3. responsibilities of the organisation.

Compliance with this framework provides a degree of assurance or confidence in the safety performance of the system. Accidents involving civil aviation aircraft are nowadays extremely rare events. Subsequently, significant operational experience is needed before estimates of the actual safety performance of the regulated system can be determined with a reasonable degree of confidence.

Differences between the Conventionally Piloted Aircraft (CPA) and UAS safety paradigms will influence how the safety risks associated with the operation of the two technologies can be most effectively managed. Researchers have explored a number of different components of the safety risk management process and its application to UAS. A number of factors that can give rise to differences in the specification of safety goals for CPA and UAS are presented in Clothier & Walker (2013). Issues specific

to the evaluation risks associated with UAS are explored in Clothier, Williams, Fulton & Lin (2013). An overview of the safety risk management process and some of the issues specific to its application to UAS can be found in Clothier & Walker (2013).

The default position of NAAs is to seek to apply and adapt the existing CPA framework of regulations to UAS (Clothier et al. 2011). It has been argued that the existing regulations may not provide for an effective management of the risks for UAS due to the differences between the CPA and UAS safety risk paradigms (Hayhurst et al. 2006, Clothier et al. 2011, Clothier & Wu 2012). One of the most significant differences between the two paradigms is that the primary risks associated with UAS operations are to people external to the system (i.e., third parties on the ground or secondary parties onboard other aircraft). Conversely, the primary risks associated with CPA operations are to people onboard the aircraft (Clothier & Walker 2013). Another difference, and the principle subject of this paper, is in the integral contribution of the "human element" to the safe operation of the system.

For UAS, many of the functions that once were performed by a human pilot are now provided by hardware and software systems. The application of the existing CPA framework of regulations and standards may not account for differences in the allocation of functions to hardware, software and "human" components of an UAS. The allocation of functions will also depend on the LoA of the UAS[1]. Hayhurst et al. (2006) states that

> "It is not clear, however, whether existing regulations that are based on a historical pairing of pilot and plane can be adapted to accommodate UASs, or whether UASs constitute a fundamentally different category of aircraft requiring their own set of regulations."

This issue is further explored in this paper. In order to satisfy the safety performance objectives established for UAS, safety assurance is required at higher 'levels' within the aviation safety system. New tools for providing assurance in the safety of autonomous aviation systems are needed. These tools must be capable of assessing the safety of the autonomous system under a wide range of operating conditions, missions and failure scenarios.

The rest of the paper discusses the current safety assurance framework and then proposed a new framework for risk management of UAS with increased levels of autonomy.

2 The Safety Assurance Framework

Figure 1 depicts a hierarchical model of the aviation safety system. This model describes different levels of organisational complexity of the aviation system. The focus of this section is on describing the different mechanisms for providing assurance in the safety at the level of *Operations* and below.

The *Operators* component describes the framework of organisational policies, procedures and resources in which *Scenarios* are conducted. A *Scenario* encompasses the interaction of the components of *Aircraft*, *Mission*, and *Weather*. While the *Airspace System* (which includes the various Classes of Airspace and the Air Traffic Management (ATM) services provided) is determined at a level higher than *Scenarios* the *Airspace Requirements*, which draws on the specifics from the *Airspace System*, is determined by the *Mission*. The organisational environment is a significant factor in the safety of aviation activities. However, further consideration of the *Operators* component is beyond the scope of this paper.

The lower tiers of the model are inspired by the *SHEL Model* (Edwards 1972). Where, *'S'* stands for the software, *'H'* for the hardware, and *'L'* for the liveware (or human) components of an *Aircraft* and *'E'* for the aircraft environment (referred to as the component of *Weather* in Figure 1). The component of *Liveware* represents the interaction of a team of *Individuals* performing a range of *Roles*. The behaviour of a *Machine* emerges from the interaction of both of its components of *Hardware* and *Software*, and the behaviour of the *Aircraft*, through the interaction of the components of *Machine* and *Liveware*, for given *Missions* and *Weather* conditions.

The existing framework of regulations for CPA and its relationship to the components of the model illustrated in Figure 1 is discussed in the following section.

2.1 Safety Assurance for CPA

Historical aviation accident and incident data can be used to provide estimates of the actual safety performance of CPA at the system levels of *Operations* and *Scenario*. The observed "safety performance" is largely due to the safety assurance provided by a framework of regulations, standards and procedures governing different components of CPA *Operations*. The regulatory framework for CPA separates regulations pertaining to the initial and continuing airworthiness of the system (e.g., requirements on the design, manufacture and maintenance of the *Machine*) from those regulations pertaining to the operation (i.e., independent of the *Mission* and *Weather*), and from the training, licensing, and proficiency of the aircrew (i.e., the component of *Liveware*). Each component of the regulation can be thought of as having a 'contribution' to the overall safety performance of CPA *Operations*. The separate contributions are illustrated in Figure 2. It is important to note that in reality the safety performance of CPA *Operations* cannot be reduced to discrete contributions. Safety performance is an emergent property, specifically, an irreducible property and one "which is not determined solely from the properties of the system's parts, but which is additionally determined from the system's structure and behaviour" (Thomé 1993). However, for the purposes of providing a simple illustration for discussion it is represented as an "aggregate" relationship.

For CPA, the primary entities at risk are the crew and passengers onboard the aircraft (Clothier & Walker 2013). Consequently, CPA regulations implicitly aim to limit or eliminate harm to those aboard the aircraft, and secondarily to those over-flown (Hayhurst et al. 2006). This philosophy is reflected in the framework of initial and ongoing airworthiness regulations, which govern the design, manufacture and maintenance of a CPA. As described by Haddon & Whittaker (2002), as far as is practicable, the airworthiness codes of regulatory requirements avoid any presumption of the purposes for which the aircraft will be used in service. Referring to Figure 1, the airworthiness regulations provide safety assurance below the level of the *Aircraft* at the component of the *Machine* (i.e., largely independent of the components of *Mission* and *Weather*). The airworthiness

[1] A review of different frameworks for describing levels of autonomy is provided by Clothier, Perez and Williams (Clothier, Perez & B. Williams 2013)

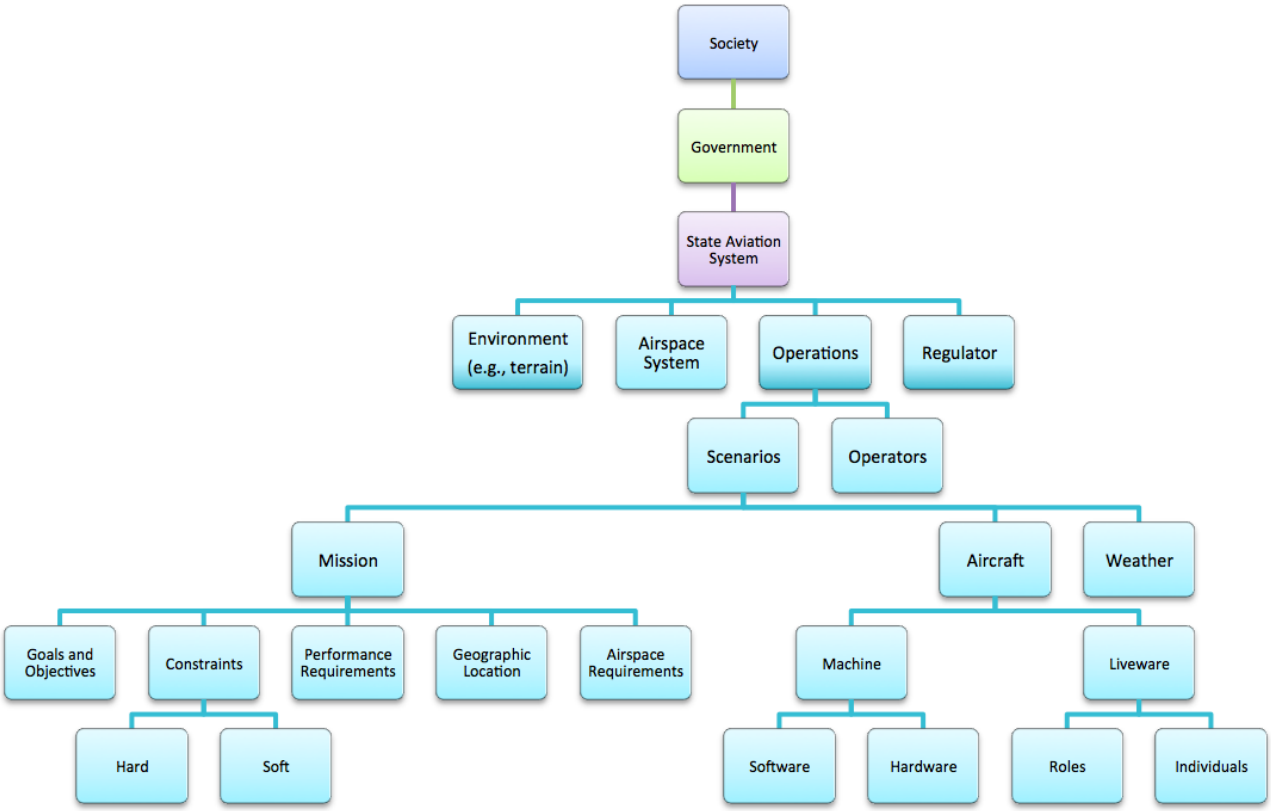

Figure 1: A model of the civil aviation regulation framework.

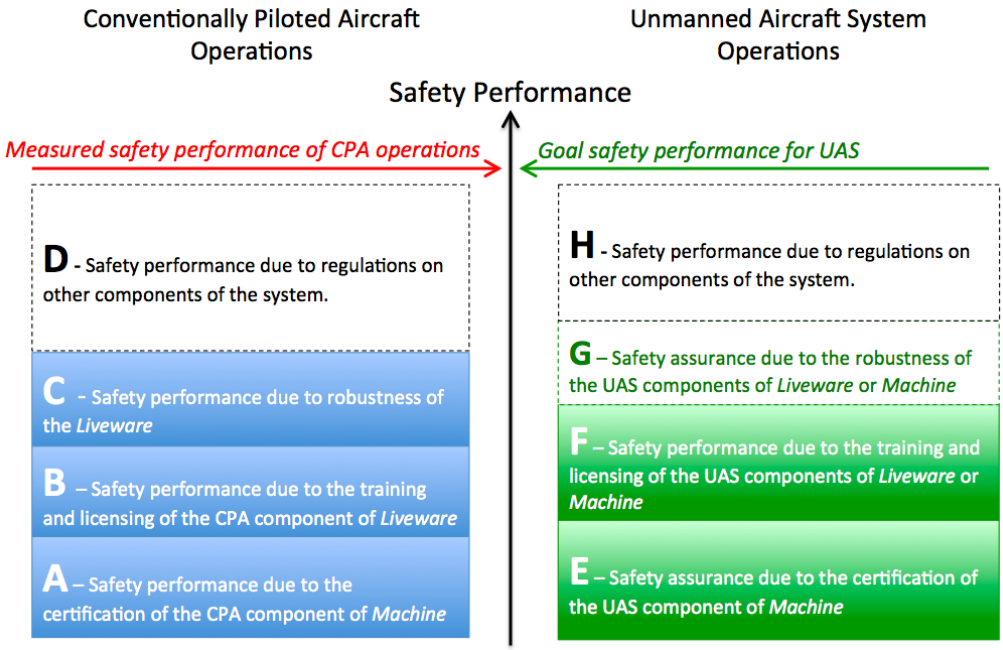

Figure 2: Comparison of Contributions to Safety Performance of CPA and UAS Operations.

regulations are specific to the component of the *Machine*. Compliance to these regulations and standards is recognised through the issuing of a Type Certificate and Certificates of Airworthiness—see, for example, CASA (2000). The airworthiness regulations and standards contribute to a proportion of the overall safety performance of CPA *Operations*, labelled as region '*A*' in Figure 2. It should be highlighted that while an *Aircraft* design (formalised by the Type Certificate) will limit the types of *Mission* and *Weather* in which operations can take place there is no certification requirement at the level of *Scenarios*. There is a consideration of suitability conducted at the *Operations* level, including continuation training requirements.

CPA airworthiness regulations are defined largely independent of the human component of an aircraft system (the component of *Liveware*). Whilst standards for human-machine interface and handling qualities are defined, the airworthiness standards are implicitly defined based on the assumption of a "nominal pilot". Separate regulations and standards pertaining to the training, licensing and currency of aircrew provide safety assurance in the component of *Liveware*. Aircrew are trained to operate safely a particular type of aircraft under a wide range of normal and abnormal operating conditions, missions and emergency scenarios. The proficiency of the crew in performing these functions is provided through training, examination and through demonstration by means of a flying test(s). Satisfactory performance results in the issuing of a licence. Training and licensing of *Liveware* contributes to a proportion of the overall safety performance of CPA *Operations*. This contribution is represented as region '*B*' in Figure 2. It is important to note that the performance of the *Liveware* can be highly varied and unpredictable. The issuing of a pilot licence provides no guarantee that all pilots will perform as trained. Subsequently, the contribution of a trained pilot to the overall safety performance of the system can vary significantly (i.e., the region '*B*' is highly dependent on the individual). Continuation training, performance and currency monitoring is undertaken at the *Operations* level.

There is an additional contribution to the overall safety performance of CPA *Operations*, which arises from the ability of the pilot (*Liveware*) to handle undesired, emergent and unforeseen scenarios that arise due to failures in, or interactions between, the subcomponents of the *Aircraft*, or due to the interaction of the *Aircraft*, *Mission* and *Weather*. This contribution can be viewed as a "Guardian Angel" function. Examples of where the *Liveware* has successfully adapted to accommodate such undesired scenarios would include Air Canada Flight 142, known as the "Gimli Glider" (Williams 2003). As the Gimli Glider example highlights, there is an additional contribution made by the *Liveware* to the overall safety performance of CPA operations. (illustrated as region '*C*' in Figure 2). It is important to note that the component of *Liveware* is not infallible nor can it handle all foreseeable scenarios. As CPA accident data would attest, the component of *Liveware* can be both a mitigator and significant contributor towards CPA accidents. Thus, the region '*C*' represents the net contribution of all pilots to the safety performance of CPA *Operations*. Hypothetically, if safety performance *could* be measured down to the level of individual pilots, then for some pilots the net contribution could be negative.

A significant contributor to the safety performance of CPA is provided by the *Liveware*. This contribution is not due to the safety assurance provided by the existing regulations and standards. This is because the regulatory framework for CPA only provides safety assurance with respect to the individual components of the *Aircraft* (e.g., the separate components of *Machine* and *Liveware*) and not with respect to the system as a "whole". The question of interest in this paper, is whether the same framework of regulations for CPA would result in the same degree of safety performance if applied to UAS. Of particular interest are those UAS which exhibit a high LoA (i.e., where the role of the human in the operation of the UAS is significantly reduced).

2.2 Application to UAS

The high-level safety objective for UAS is that they demonstrate, as a minimum, a level of safety performance equivalent to that currently demonstrated by CPA. This objective is commonly referred to as Equivalent Level of Safety (ELoS). A summary of qualitative and quantitative statements of the ELoS objective is discussed by Clothier & Walker (2013). We can establish the "goal" safety performance for UAS with reference to the "measured or observed" safety performance for CPA, as illustrated in Figure 2. The primary risks associated with the operation of CPA are to those onboard the aircraft (first parties). Whereas for UAS, the primary risks are to those people onboard other aircraft (secondary parties) or to those on the ground (third parties) (Clothier & Walker 2013). The ELoS objective for UAS must therefore be determined in relation to the risks CPA pose to second and third parties.

With reference to the hierarchical model presented in Figure 1, the ELoS objective can be defined at the component of the UAS *Operation*. Regulations on other components of the aviation safety system (e.g., *Operators*) will contribute to the overall safety performance of the system (illustrated by Regions '*D*' and '*H*' in Figure 2). Discussion in this paper is focusses on the safety-contributions provided by the regulations on the components of the *Aircraft*. Thus, the comparison being made in this section assumes the components of *Mission* and *Weather* are constant.[2]

The application of the existing framework of regulations and standards governing the airworthiness of UAS will have a hypothetical contribution to the overall safety performance exhibited by UAS *Operations* (illustrated as region'*E*' in Figure 2). Similarly, training and licensing of the human RP will also provide a degree of assurance in the safety performance of UAS operations (illustrated as region '*F*'). Similar to the pilot of a CPA, the presence of a human RP can also provide a degree of "robustness" to the UAS operation, which contributes to the overall safety performance of the UAS (illustrated as region '*G*'). These contributions ('*E*', '*F*' and '*G*') will depend on the LoA of the UAS. The higher the LoA of the UAS, the less the human RP is involved in the operation of the UAS and the more the UAS subsumes the role (i.e., functions) of the RP (Clothier & Walker 2013). It follows that as the LoA increases, the more functions are required to be performed by the *Machine*.

[2]It is important to note that applying CPA regulations to UAS (at the level of the *Machine*) will not account for potential differences in the nature of the typical *Missions* performed by UAS or differences in the typical environments (*Weather*) in which UAS are operated. For example, High Altitude Long Endurance UAS or UAS operating at low levels in urban environments. (Clothier & Wu 2012) illustrates how the application of existing CPA system reliability requirements to UAS can, in some cases, lead to an unacceptable level of risk at the level of an UAS *Operation*.

Consider an UAS with a LoA where the RP performs the exact same functions as a human pilot of a CPA, with the only difference being that the RP is remotely located from the Unmanned Aircraft (UA). Despite the 'functions' performed by the RP and conventional pilot being the same, their relative contribution to the safety performance of their respective operations will not be the same. This difference in safety performance is a result of the limited experience in the operation of UAS in civil airspace and as a consequence the training and licensing of RPs may not provide the same degree of safety assurance as that typically provided by the training and licensing of pilots of CPA. There are also human performance considerations unique to UAS (Fothergill et al. 2013) that will impact on the net contribution of the *Liveware* to the overall safety performance of an UAS *Operation*. For example, psychological differences (e.g., trust in autonomy, the lack of "pilot shared fate" with the UA), and differences in RP situational awareness, etc. In sum, even if UAS exhibited the same degree of reliability in the *Machine* and the functions performed by the *Liveware* were the same, the relative contributions and the overall safety performance exhibited is likely to be different for UAS compared to CPA.

Now consider the other extreme, where an UAS has a LoA where the human RP no longer has a role in the flight of the UAS. All of the functions of the *Liveware* are now provided by the component of *Machine*. The safety-contributions previously provided by the *Liveware* (i.e., 'F' and 'G') must now be provided by the *Machine*. Existing software and hardware standard can provide a degree of safety assurance in the *Machine* for those functions for which RP are trained (i.e., 'F'). However, the existing safety contribution due to the adaptability of the RP (i.e., region 'G') would not be assured. Specifically, can the functions previously provided by *Liveware* and now provided by the *Machine* still be adequately performed in the presence of failures in the *Machine* or under unforeseen conditions?

To answer this question, safety assurance must be provided at the component-level of the UAS *Scenarios* in addition to the existing safety assurance mechanisms for the components of *Software* and *Hardware*. New certification tools are needed that can assess the UAS *Scenarios* as a "whole" under varying missions, environmental conditions and failures. The next section (Section 3) describes one possible tool that can be used to provide such a certification assessment.

2.3 Implications for CPA

The discussion of Safety Assurance for CPA reflects current practice. It is important to note that the discussion based on UAS are not limited in applicability to UAS, and it is highly likely that as safety assurance is studied from a UAS context that it will influence and update thinking and practices in CPA.

3 A Proposed Framework for UAS Robust Autonomy Certification

The management of risk of UAS requires quantifying both uncertainty and consequences of outcomes, where the latter refers to behaviour of the system as a whole. While CPA operations may be conducted under similar uncertainty, the previously discussed abilities of the *Liveware* has been the basis for acceptable safety and performance. In UAS, these decisions under uncertainty have shifted from *Liveware* to *Machine*, or some combination of both. The uncertainty is related to three characteristics:

- the actual environmental conditions that the UAS will encounter during the missions (weather and complexity of the operational airspace),
- the reliability of the different components and subsystems of the UAS (airworthiness), and
- the ability of the autonomy to make the rational decisions regarding guidance, navigation and motion control in both normal and anomalous conditions (robust decision making).

All these characteristics are encapsulated in the term *Robust Autonomy*, which describes the ability of an autonomous system to either continue or abort its operation whilst not breaching a minimum level of acceptable safety in the presence of anomalous conditions (Perez et al. 2011*b*, Perez & Williams 2012, Perez et al. 2011*b*). Robust autonomy encompasses both safety and reliability. In addition, the qualifier "*robust*" highlights the feature that the autonomous operation is being considered under both normal and anomalous conditions.

In the absence of a regulatory framework for the certification of highly autonomous UAS, we propose a process based on six steps:

1. Adopt a set of missions for which a UAS is being certified.
2. Adopt the relevant measures of safety and performance and their associated sets of acceptable level.
3. Adopt the envelope of operational conditions within which the missions must be performed.
4. Conduct an evaluation of autonomy to compute probabilities of maintaining acceptable safety and performance.
5. Present a certification case to the NAA.
6. NAA makes a certification decision according to the probabilities and levels of risk deemed acceptable.

The outcomes of Steps 1 to 3 should be a clear set of requirements for particular missions and classes of UAS deemed appropriate for these missions. For example, UAS operations for bush fire monitoring, sea search and rescue, and traffic monitoring over populated areas may all have different requirements and envelopes of weather and environments in which they are allowed to operate. In addition the characteristics of the mission may also have a bearing on the required sensor and actuator physical redundancy. Step 4 requires a probabilistic assessment of safety and performance, which takes into account the reliability of the *Machine* (e.g., the airframe, sensors, and actuators), the likelihood of the operational conditions (e.g., weather, faults, failures, and complexity of the space in which the mission is conducted), and the capabilities of the autonomy to make rational decisions in regard to mission execution, guidance, navigation, communication, and motion control (the equivalent of the *Liveware* in the CPA). Step 4 is expected to be conducted by third-party testing organisation using methods and tools approved by the NAA. Step 5 involves the operator seeking certification to gather information to put a case forward for the NAA. This

will include a certificate from third-party testing organisation. Step 6 involves the decision making by the NAA about certification, which requires the quantification of the consequences of the potential outcomes as perceived by the NAA.

The development of steps 1 to 5 above are to be addressed in collaboration with the NAA. In the following, we discuss specific aspects of each of the steps of the proposed framework above.

3.1 Step 1 - Classification of UAS Missions

UAS are designed for specific operations and environments under which the operations need to be conducted. The objective of Step 1 is to agree with the NAA on a classification of UAS type and its associates class of missions. The class may be related to operations such as search and rescue at sea, bush fire monitoring, border patrol, mineral exploration, *etc.* The classes can be defined by attributes such as mass, sensor redundancy, actuator redundancy, limits on the envelope of environmental conditions for which operations can be conducted, whether operations are to be conducted over populous or non-populous areas, and whether the operations are to be conducted in operational spaces of limited complexity (segregated airspace). Once the classes are defined, the measures of safety and performance associated with each class needs to be determined in Step 2.

3.2 Step 2 - Safety and Performance Indicators

UAS within a particular class should be able to conduct specific operations or missions in prescribed environments. The measures of safety and performance can be evaluated in terms of performance indices r_i $(i = 1, 2, ..., l)$ related to safety requirements of the operation and perhaps also performance attributes that can impact on safety. For example, indices related to safety include separation between aircraft, location with respect to no-fly zones, ability to declare and communicate emergencies, ability to re-route after declaring emergencies, ability to detect and accommodate certain faults, detect and avoid, kinetic energy on emergency landing, *etc.* Performance attributes with bearing on safety are related to the particular aircraft flight envelope. Hence, climb rate, bank angle, loading factor, airspeed, and angle of attack, may be considered in relation to a flight envelope. The adage "Aviate, Navigate, Communicate" can assist in identifying measures of safety and performance, as well as providing a prioritised ordering.

For each quantifiable performance index r_i $(i = 1, 2, ..., l)$, we can associate a set $\mathcal{R}_i$, such that satisfactory performance is attained whenever the value of the index is in the set for the complete mission, namely $r_i \in \mathcal{R}_i$. Then, we can define the *Event of Satisfactory Performance* as such in which a performance index remains inside its region of satisfactory performance for the complete mission:

$$S_i = \{r_i \in \mathcal{R}_i\}, \quad i = 1, 2, ..., l. \tag{1}$$

These events are a logical statement or hypotheses which once a mission is tested can either be true or false.

3.3 Step 3 - Envelope of Operational Conditions

The missions for which the aircraft is being certified are to be conducted under an envelope of operational conditions which encompass weather, aircraft health, and complexity of the airspace. The uncertainty as to which weather condition W_j $(j = 1, 2, ..., m)$ can occur during the mission can be described by the probability distribution $P(W_j|I)$, where I represents background information. These probabilities can be estimated from meteorological data for a particular geographical location and time of the year. Note that the weather conditions to be considered for the operation of the UAS may depend on the type of mission. For example, an UAS used for bush fire monitoring is expected to operate in potentially turbulent high windconditions, whereas a UAS used for aerial spraying of crops may be certified only for operation in light wind conditions.

The UAS may also be subjected to faults, $F_k (k = 0, 1, ..., n)$, which can be associated with the aircraft's actuators, sensors, communication link, changes in dynamics, and the presence of other aircraft in the operational airspace. The condition F_0 denotes the faultless or nominal case (healthy platform operating in anticipated airspace complexity). The uncertainty as to which fault may occur during the mission is described by the probability distribution $P(F_k|I)$. If a fault is associated with a component or a subsystem, for example a servo for a control surface, then $P(F_k|I)$ can be computed from the failure rate function of the component or system (Singpurwalla 2006, Hamada et al. 2010). Faults associated with the complexity of the operational space, $P(F_k|I)$ can be computed from air traffic data or other background information.

3.4 Step 4 - Probabilistic Assessment of Safety and Performance

The work by Perez et al. (2011*b*,*a*) provides a probabilistic assessment of robust autonomy in terms of behaviours. This is motivated by the fact that when pilots are evaluated, it is their behaviour in specific situations which is being assessed rather than their neurophysiological process that leads to the behaviour. A similar procedure can be followed for assessing the rationality of decision making of the autonomy, in which the evaluation is done without specific knowledge of the implementation of autonomous decision making. The potential to evaluate the performance of the autonomy of the *Machine* under failure is critical to providing assurance in the robustness of the UAS (i.e., the additional contribution '*G*' in Figure 2.)

The performance during the mission can then be assessed in terms of the predicted probabilities of the events of satisfactory performance S_i as defined in Step 2 (Perez et al. 2011*a*). The hardware that implements the autonomous decisions can be connected to a hardware-in-the-loop (HIL) simulator and tested under the selected set of weather and operational conditions (Step 3). The data, D, collected during a HIL test consists of aircraft motion, location, and also information exchanged over available communication channels. These data can be used to compute the predicted distribution $P(S_i|D, I)$. These probabilities are called *Measures of Robust Autonomy* (Perez et al. 2011*b*) and can be computed via marginalisation:

$$P(S_i|D, I) = \sum_j \sum_k P(S_i|W_j, F_k, D)\, P(W_j|I) P(F_k|I). \tag{2}$$

Each of these measures involves different aspects of the system which contributes to its total reliability and in turn the overall safety performance of

the UAS operation. The distributions $P(W_j|I)$ and $P(F_k|I)$ capture uncertainty about the environment in which the system is to operate. The distribution $P(S_i|W_j, F_k, D), (i = 1 : l)$ evaluates the quality of autonomous decision making of the UAS under a particular scenario given by the combination W_j, F_k. The latter encompasses aspects of robustness and performance of the vehicle control system, fault detection and diagnosis system, and on-line decisions about reconfiguration of the control system and mission rerouting and trajectory planning.

The distribution $P(S_i|W_j, F_k, D)$ in (2) is related to the concept of *coverage* discussed by Wu (2004), that is, the probability of maintaining a desired level of performance and safety given that a particular scenario has occurred. In the context of this paper, coverage encompasses not only the low-level motion controller but, depending on the degree of autonomy of the platform, also the guidance, communications and navigation systems.

Note that the distribution $P(S_i|W_j, F_k, D)$ is related to the level *Mission* in Figure 1 since the risk measures or performance indices are defined for a set of operations or missions as discussed in Step 2. Also, the weather $P(W_j|I)$ and fault $P(F_k|I)$ distributions are related to the levels of *Aircraft* and *Weather* in Figure 1. Therefore, the measures of robust autonomy $P(S_i|D, I)$ in (2) relate to the level *Scenarios* in Figure 1.

Since the events S_i, defined in (1), are proper hypotheses about the system's behaviour, which can either be true or false, the probability of obtaining a number of successes in a certain number of missions can be modelled using the Bernoulli distribution where π_i is a parameter that that gives the probability of success in one mission. If we collect data D from N replications of evaluation of performance using a hardware-in-the-loop simulator, we can update our prior $p(\pi|I)$ to the posterior

$$p(\pi_i|D, I_{jk}) = \frac{p(D|\pi_i, I_{jk})p(\pi_i|I)}{\int p(D|\pi_i, I_{jk})p(\pi_i|I)\, d\pi_i}, \qquad (3)$$

where, $I_{jk} = \{W_j, F_k, I\}$ represents the information related to the particular condition being tested. The posterior densities $p(\pi_i|D, I_{jk})$ encode our uncertainty on the value of π_i under the condition W_j, F_k.

From the posterior $p(\pi_i|D, I_{jk})$, we can obtain the coverage probabilities $P(S_i|D, W_k, F_k, I)$ as predicted probabilities for the success in a single mission (Perez et al. 2011*a*):

$$P(S_i|D, W_k, F_k, I) = \int_0^1 \pi_i\, p(\pi_i|D, I_{jk})\, d\pi_i. \qquad (4)$$

In some cases, it may be convenient to have a single figure of merit for robust autonomy. The natural procedure to obtain this figure would be to evaluate the probability that all the indices are jointly within their regions of acceptable performance, namely,

$$\begin{aligned} &P(S_1, \ldots, S_l|D, I) \qquad (5)\\ &= \sum_j \sum_k P(S_1, \ldots, S_l, W_j, F_k|D, I)\\ &= \sum_j \sum_k P(S_1, \ldots, S_l|W_j, F_k, D)\, P(W_j|I)P(F_k|I). \end{aligned}$$

Details of these computations including a discussion on the choice of prior in (3) according to the principle of maximum entropy can be found in (Perez et al. 2011*a*).

If some indices are not within their region of acceptable performance then it would be expected that they degrade in a manner that meets the priorities established in the adage "Aviate, Navigate, Communicate".

3.5 Step 5 - Present Certification case

The NAA is not likely to conduct Step 4. The assessment can be conducted by an NAA-approved third-party testing organisation. The testing organisation would issue a report on the outcomes of the assessment, which the UAS *Operator* could present to the NAA as part of a case for certification.

Inspiration for the framework proposed in this paper is taken from the offshore shipping industry, where a third-party industry dedicated to assessment has assisted the Classification Societies[3] to develop regulations for testing of behaviours of on-board power management systems and ship motion control systems using hardware in the loop. The services of the assessment industry can also be used by UAS developers to obtain feedback on aspects of autonomy that should be improved. For example, (Perez et al. 2011*b*) discusses a case where information about such an evaluation identifies the need to either improve on fault detection and handling and mission re-planning or reduce the probability of failure of an actuator. From the point of view of design, the framework can suggest areas that may need improvement (for example fault-tolerance, mission re-planning and guidance, or increase the reliability of particular component or sensor to reduce its failure probability), the actual solution to achieve such improvement is not, however, the objective of proposed framework.

3.6 Step 6 - The Certification Decision

Given a case presented for certification, the NAA must make a decision. This decision must be made under uncertainty since the information provided by Step 4 is a probability that the system will attain the required levels of safety and performance, namely (5), mandated by the class under which the UAS seeks certification. Note that this is similar to the decision made as to issuing a licence to a pilot - a decision under uncertainty.

The NAA is likely to compare the measures of robust autonomy with threshold probabilities. To aid in the setting of these thresholds to make such a decision, we can look at normative decision theory (Peterson 2009, Singpurwalla 2006, Berger 1985, Jaynes 2003). In this theory, a decision problem has three key ingredients:

$$\Pi = \langle \mathcal{A}, \mathcal{X}, \mathcal{O} \rangle,$$

where

- $\mathcal{A} = \{A_1, A_2, \ldots, A_m\}$ is the set of actions being considered,
- $\mathcal{X} = \{X_1, X_2, \ldots, X_n\}$ is the set of states of nature about which there is uncertainty: $P_k = P(X_k)$,
- $\mathcal{O} = \{O_{ij}\}, i = 1, 2, \ldots, m, j = 1, 2, \ldots, n$ is the set outcomes.

[3] The Classification Societies are entities that develop standards and issue certificates for design and construction practices as well as operations of ships and offshore structures. These certificates are required by insurance companies for insuring vessels.

To solve a decision problem, we require a *Loss function*, which quantifies the consequences of the outcomes. That is the loss function L maps the states of nature and the actions into numbers that measure the consequences of the outcomes:

$$L : A_i, X_j \mapsto L_{ij}(O_{ij}).$$

That is, L_{ij} measures the consequence of taking the action A_i were X_j be the true state of nature.

A decision can then be made by adopting a *Decision Criterion*, which selects the preferred action based on the consequences of the outcomes and the uncertainty of the states of nature:

$$C : \{L_{ij}, P_k\} \mapsto A^\star.$$

For the decision about certification, a simplified decision problem involves a set of a binary actions that the NAA can make:

$$A_1 \text{ - Certify}, \qquad A_2 \text{ - Do not certify.}$$

The states of nature of interest can be defined as the state in which all the performance indices are jointly satisfied and its complement:

$$X_1 = S_1 \cap S_2 \cap \cdots \cap S_l, \qquad X_2 = \bar{X}_1,$$

and then

$$P_1 = P(X_1) = P(S_1, \ldots, S_l | D, I),$$
$$P_2 = P(X_2) = 1 - P_1.$$

The outcomes are

- O_{11} - Certifying a UAS that will satisfy the safety and performance requirements.
- O_{12} - Certifying a UAS that will not satisfy the safety and performance requirements.
- O_{21} - Not certifying a UAS that will satisfy the safety and performance requirements.
- O_{22} - Not certifying a UAS that will not satisfy the safety and performance requirements.

The loss function in terms of the parameters L_{ij} shown in Table 1. A positive value L_{ij} represents

Table 1: Decision Matrix

	X_1- Safe	X_2-Not safe
A_1- Certify	L_{11}	L_{12}
A_2- Do not certify	L_{21}	L_{22}

a loss, and a negative value represents a gain or a reward. For example,

- L_{11} is a negative number, which reflects the reward for certifying a UAS that satisfies the safety and performance requirements.
- L_{12} is a positive number, which reflects a loss for certifying a UAS that is not safe.
- L_{21} is a positive number, which reflects a loss for denying certification to a safe UAS. (L_{21} is likely to be less than L_{12} as it would be a worse to certify an unreliable UAS compared to denying certification to a worthy UAS.)
- L_{22} is a negative number, which reflects the reward for denying certification to an unsafe UAS.

These four numbers and the scale in which they are measured reflect the the attitude of the NAA (risk proneness or risk averseness). The determination of the these numbers is a complex multidisciplinary task that will require involving not only the NAA, but also industry, social scientists and other subject matter specialists.

At the time of making a decision, whether the true state of nature is X_1 or X_2 will not be known with certainty: only their probabilities P_1 and P_2 will be known through the calculations made in Step 4. The decision criterion that satisfies the requirements of consistency and rationality is that of taking the action that minimises the Bayesian risk (Jaynes 2003). The Bayesian risk is the expected loss over the posterior distribution of the uncertain states of nature at the time of making the decision (Singpurwalla 2006, Berger 1985), namely,

$$A^\star = \arg\min_{A \in \mathcal{A}} E^X[L(A, X)],$$

where $E^X[\cdot]$ denotes the expectation operator with respect the distribution of X. That is, the NAA should take the action that gives the minimum of the two risks:

$$\rho(A_1) = L_{11}\, P_1 + L_{12}\, P_2,$$
$$\rho(A_2) = L_{21}\, P_1 + L_{22}\, P_2.$$

As discussed by Peterson (2009), the choice of expected loss as decision criterion is not to be interpreted as making the decision that minimises the loss in average (as if decisions were made several times). This criterion arises from the satisfaction of axioms of rationality and consistency, from which it follows that rational agents making decisions behave *as if* they minimise the expected loss.

4 Conclusion

This paper discusses a model of the civil aviation safety framework and shows how the current assessment of reliability and risk for conventionally piloted aircraft may not provide an appropriate framework for the same degree of assurance in the safety of Unmanned Aircraft Systems (UAS) operations of varying levels of autonomy. This is because existing approach does not certify the system as a whole. A new framework for certifying UAS, based on the principles of risk management is proposed. This framework arises from combining quantified measures of risk with a probabilistic assessment and normative decision making. The decision making requires quantifying probabilities associated with the measures of risk and also consequences of outcomes related to the behaviour of an UAS. These probabilities are measures of uncertainty about the events of satisfactory safety and performance, and they are computed using a Bayesian approach from an assessment under both nominal and anomalous scenarios described by faults. The framework poses the decision making as a normative decision problem and solves it in terms of the minimisation of the expected loss - a criterion which satisfies the requirements of consistency and rationality of probability theory.

5 Acknowledgment

This research was in part supported under the Robust Autonomous Systems collaboration between the University of Newcastle and Boeing.

References

Berger, J. (1985), *Statistical Decision Theory and Bayesian Analysis.*, Springer.

CASA (2000), *Aircraft Airworthiness Certification Categories and Designations Explained - Advisory Circular, AC 21.1(1)*, Civil Aviation Safety Authority (CASA), Canberra, Australia.

Clothier, R., Palmer, J., Walker, R. & Fulton, N. (2011), 'Definition of an airworthiness certification framework for civil unmanned aircraft systems,', *Safety Science* **49**(6), 871–885.

Clothier, R., Perez, T. & B. Williams, B. (2013), A review of the concept of autonomy in the context of the safety regulation of civil unmanned aircraft systems, *in* 'Australian System Safety Conference (ASSC2013), Adelaide, Australia.'.

Clothier, R. & Walker, R. (2013), *Safety Risk Management of Unmanned Aircraft Systems*, Springer Science + Business Media B.V., Dordrecht, Netherlands., chapter 3.

Clothier, R., Williams, B., Fulton, N. & Lin, X. (2013), ALARP and the risk management of civil unmanned aircraft systems, *in* 'Australian System Safety Conference (ASSC2013), Adelaide, Australia.'.

Clothier, R. & Wu, P. (2012), A review of system safety failure probability objectives for unmanned aircraft systems, *in* 'Eleventh Probabilistic Safety Assessment and Management Conference (PSAM11) and the Annual European Safety and Reliability Conference (ESREL 2012)', Helsinki, Finland.

Edwards, E. (1972), Man and machine: Systems for safety, *in* 'British Airline Pilots Associations Technical Symposium', British Airline Pilots Associations, London.

Fothergill, S., Clothier, R. & Coyne, J. (2013), Human performance considerations for civil unmanned aircraft systems, *in* 'Australian System Safety Conference (ASSC2013)', Adelaide, Australia.

Haddon, D. & Whittaker, C. (2002), *Aircraft Airworthiness Standards for Civil UAVs*, UK Civil Aviation Authority (CAA), London, UK.

Hamada, M., Wilson, A., Reese, C. & Martz, H. (2010), *Bayesian Reliability*, Springer.

Hayhurst, K., Maddalon, J., Miner, P., DeWalt, M. & McCormick, F. (2006), Unmanned aircraft hazards and their implications for regulation, *in* 'IEEE/AIAA 25th Digital Avionics Systems Conference (DASC)', Portland, OR, United States of America. Portland, OR, USA.

Jaynes, E. (2003), *Probability Theory - The Logic of Science*, Cambridge University Press.

Perez, T. & Williams, B. (2012), Assessment of robust autonomy for unmanned systems – progress and challenges., *in* 'AUVSI's Unmanned Systems North America', Las Vegas, NE, USA.

Perez, T., Williams, B. & de Lamberterie, P. (2011*a*), Computational aspects of probabilistic assessment of uas robust autonomy, *in* '28th International Congress of the Aeronautical Sciences ICAS, Brisbane, Australia.'.

Perez, T., Williams, B. & de Lamberterie, P. (2011*b*), Evaluation of robust autonomy and implications on uas certification and design, *in* '28th International Congress of the Aeronautical Sciences ICAS, Brisbane, Australia.'.

Peterson, M. (2009), *Introduction to Decision Theory*, Cambridge introduction to phylosophy, Cambridge Unversity Press.

Singpurwalla, N. (2006), *Reliability and Risk*, Wiley Series in Probability and Statistics.

Thomé, B. (1993), *Systems Engineering: Principles and Practice of Computer-Based Systems Engineering*, John Wiley & Sons Ltd, New York.

Williams, M. (2003), 'The 156-tonne gimli glider', *Flight Safety Australia* **27**, 22–27.

Wu, N. (2004), 'Coverage in fault-tolerant control', *Automatica* **40**(4), 537–548.