EVALUATION OF ROBUST AUTONOMY AND IMPLICATIONS ON UAS CERTIFICATION AND DESIGN

Tristan Perez*,*, Brendan Williams**, Pierre de Lamberterie***
***School of Engineering, The University of Newcastle, Callaghan NSW 2308, AUSTRALIA,**
****Boeing Research & Technology Australia, GPO Box 767 QLD-4001, AUSTRALIA,**
***** CeSOS, Norwegian University of Science and Technology, Trondheim, NORWAY**
tristan.perez@newcastle.edu.au; brendan.p.williams@boeing.com;
pierre.delamberterie@uon.edu.au

Keywords: *Robust Autonomy, UAS design & certification*

Abstract

As the number of potential applications of Unmanned Aircraft Systems (UAS) grows in civilian operations and national security, National Airworthiness Authorities are under increasing pressure to provide a path for certification and allow UAS integration into the national airspace. The success of this integration depends on developments in improved UAS reliability and safety, regulations for certification, and technologies for operational performance and safety assessment. This paper focusses on the latter and describes the use of a framework for evaluating robust autonomy of UAS, namely, the autonomous system's ability to either continue operation in the presence of faults or safely shut down. The paper draws parallels between the proposed evaluation framework and the evaluation of pilots during the licensing process. It also discusses how the data from the proposed evaluation can be uses as an aid for decision making in certification and UAS designs.

1 Introduction

The large potential for applications of Unmanned Aircraft Systems (UAS) in civilian operations and national security is putting an increasing amount of pressure on National Airworthiness Authorities (NAAs) to provide a path for certification and allow UAS integration into the national airspace [1]. The International Civil Aviation Organisation (ICAO) guidance and discussions on Unmanned Aircraft Systems (UAS) has focused on the subset of UAS termed Remotely Piloted Aircraft Systems (RPAS). This has been done to ensure adherence to ICAO's scope of authority, as well as to reduce the UAS problem space and complexity to a manageable level. The RPAS concept maintains human interaction and supervision as a key risk mitigation factor for "unmanned" operations. This mitigation is based on the expectation that the remote pilot will provide decision making and input during abnormal or emergency situations to ensure acceptable levels of safety are maintained or, in the extreme cases, to ensure minimal consequence of unfavourable outcomes. The ICAO position is likely to be followed by individual NAAs as they work towards implementing acceptable UAS regulations and guidance in the near term.

The UAS industry and research community view this as a stepping stone towards the acceptance of UAS operations with much higher degrees of autonomy than RPAS. In order for the NAAs to regulate and accept UAS with higher degrees of autonomy a viable and reliable method of certifying autonomy needs to be developed and agreed upon.

When higher degrees of autonomy of UAS are discussed in the literature, terms such as machine learning, artificial intelligence, intelligent agents, neural networks, fuzzy logic, decision trees, and Bayesian analysis are used to qualify the UAS autonomous decision making in guidance, navigation, communications and control (GNCC) [2, 3]. UAS robust performance need to be assessed and certified independently of the underlying principles by which autonomous decision making is made in the GNCC. Such performance evaluation must therefore relate only to measures of system reliability in relation to mission requirements and safety as well as airframe capabilities.

In a previous work, the authors took a first step in describing a framework for performance evaluation [4, 5]. The focus of this paper is on the parallel between the proposed framework and the evaluation of pilots during the licensing process, and how the data from the proposed evaluation can be used as an aid for decision making in certification and UAS designs. The framework discussed in this paper is a refinement of the proposal in [4, 5]—in this paper, we revert to more standard measures of reliability.

2 Robust Autonomy of UAS and its Evaluation

Robustness is a term that describes the feature of persistence of some attribute of an item in the presence of uncertain conditions [6]. In this paper, we adopt the following definition of robust autonomy [4, 5]:

Definition: *Robust autonomy describes the ability of an autonomous system to either continue its operation in the presence of faults or safely shut down.*

Robust autonomy encapsulates the reliability of the UAS physical platform and components plus the ability of autonomous decision making—to varying degrees—in relation to guidance, navigation, communications and control.

In a manned aircraft, the decision making aspect is the function of the pilot. By examining the way in which a pilot is assessed to be issued with a license, we can identify a framework that can be used to certify robust autonomy. Prior to an assessment of competency, potential human pilots are subjected to an extensive training program where performance expectations are established and procedural responses are reinforced. During an assessment for a pilot license, the testing pilot is looking for a number of key characteristics of the pilot being assessed. These characteristics include

- The ability to make decisions in a highly dynamic and potentially deteriorating environment.
- The ability to manoeuvre the aircraft in a manner that maximises the likelihood of a successful outcome without increasing risk by operating outside of the design performance envelopes of the aircraft.
- Following agreed and taught "best practice" procedures and manoeuvres.
- Maintaining the air vehicle within accepted and stated parameters (altitude, airspeed, navigation tolerances, airspace boundary compliance, and engine specifications).
- Compliance to applicable rules and regulations.

It is important to highlight from the outset that when assessing humans for the issuing of a pilot license, we do not expect that each and every emergency situation will result in a satisfactory outcome. What is assessed is that the actions of the candidate pilot conform to accepted "best practice" and that the actions position the aircraft in a manner that maximises the probability of a satisfactory outcome over an envelope of scenarios that simulate potential emergency situations. When assessing autonomy of UAS, we need to keep this in mind as the assessment cannot guarantee a positive outcome for each and every emergency scenario. The above suggests that the assessment of robust autonomy, as well as the assessment of pilots, involves making decisions under uncertainty. This calls for a probabilistic framework, which is the approached proposed in [5, 4] and the refinement discussed in this paper.

This proposed framework for certifying autonomy parallels the assessment made for human pilots. Namely, it establishes clear expectations in terms of measures of performance. These measures can be grouped according to the aerospace adage of "Aviate, Navigate, Communicate":

- Keeping inside the flight envelope (V-N diagram),
- Separation compliance (sense and avoid),
- Ability to land (recovery in normal, abnormal and emergency);
- Navigation (required location and time at location),
- Airspace compliance,
- Communications (data link and air traffic).

The above measures of performance have been listed in descending order of priority, with an expectation that the breaching of a lower priority performance measure must occur before breaching a higher order one. It is possible that the specific performance measures will be dependent on the phase of flight and whether the current flight condition is assessed as normal, abnormal or emergency.

The performance of a UAS system can be assessed in essentially the same way as pilots are assessed. Within this framework the pilot being tested is replaced by proprietary hardware, and associated software, that implements the functions of decision making in relation to GNCC. These functions can include fault detection & diagnosis and reconfiguration of the GNCC systems to accommodate faults and prevent system failure. The GNCC hardware is connected to a hardware-in-the-loop (HIL) simulation environment which can simulate different aspects of the mission under various environmental conditions and fault scenarios (sensors, actuators, UAS platform aerodynamics, other aircraft, etc.) The data collected from these HIL simulations can then be used to assess performance and therefore decision making about certification and system design. In the sequel, we discuss the main aspects of the proposed framework—the details of the computational aspects are discussed in the companion paper [9].

2.1 Performance Indices and Environment

UAS are specifically designed for particular missions and environments under which the missions need to be conducted. The measures of performance discussed in the previous section can be evaluated in terms of specific performance indices related to mission requirements and airframe capabilities. For example Table 3 shows some of the performance indices that can be adopted according to the measures of performance related to Aviate, Navigate, communicate.

Table 1 Example of performance indices for UAS Missions.

Index	Description
r_1	Climbing rate
r_2	Bank angle
r_3	Loading factor
r_4	Angle of attack
r_5	Sideslip angle
r_6	Air speed
r_7	Sense & avoid
r_8	Ability to land
r_9	Kinetic energy in emergency landing
r_{10}	Required location
r_{11}	Remain outside of a no-fly zone

For each index r_i $(i = 1,2,\ldots,l)$ we can associate a set $\mathcal{R}_i$, such that satisfactory performance is attained whenever the value of the index is in the set $\mathcal{R}_i$ for the complete mission.

The mission is to be performed under an envelope of operational conditions which encompass weather and platform faults. The weather conditions $W_j\,(j = 1,2,\ldots,m)$ refer to conditions such as: mean wind velocity, turbulence, and visibility. The uncertainty as to which weather condition can occur during the mission is described by the probability $P(W_j|I)$, 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 mission. For example, a UAS used for bush fire monitoring is expected to operate in high speed and highly turbulent winds, whereas a UAS used for aerial photography is expected to operate in light wind conditions.

The UAS platform may also be subjected to faults, $F_k \, (k = 0, 1, \ldots, n)$, which can be associates with actuators, sensors, communication link, changes in aerodynamics, and the presence of other aircraft. The condition F_0 denotes the faultless or nominal case. The uncertainty as to which fault may occur during the mission is described by the probability $P(F_k|I)$, where I represents background information. If the fault is associated with a component or a subsystem, for example a servo of a control surface, then $P(F_k|I)$ is can be taken as the reliability of the component or subsystem. That is, if a mission is to be conducted from the time t to the time $t + \Delta t$ (where the scale is related to the time where the component was first put in service) then, $P(F_k|I) = P(t < T \leq t + \Delta t | T > t, I)$, where T is the component or subsystem time to failure. That is $P(F_k|I)$ is the probability that the fault F_k will occur during the mission given that the fault has not occurred at the time of starting the mission. This probability is standard measure in reliability, and it can be computed from the failure rate function of the component or system [7].

2.2 Evaluating Performance

For each performance index, we can define the event of satisfactory performance as that in which a performance index remains inside its region of satisfactory performance for the complete mission:

$$S_i \equiv \{r_i \in \mathcal{R}_i\}. \qquad (1)$$

Note, S_i is an event that can be either true or false after a mission is evaluated.

The evaluation of the performance during the mission can be assessed in terms of the predicted probabilities of satisfactory performance each index in one mission given what we have learned from the data D related to the evaluation of the system. These probabilities can be computed by magninalising over the environmental conditions (weather and faults):

$$\begin{aligned} P(S_i|D,I) &= \sum_j \sum_k P(S_i, W_j, F_k|D,I) \qquad (2) \\ &= \sum_j \sum_k P(S_i|W_j, F_k, D) \, P(W_j|I) P(F_k|I). \end{aligned}$$

These probabilities are called *Measures of Robust Autonomy*. Each of these measures involves different aspects of the system which contribute to its reliability:

- $P(W_j|I)$ and $P(F_k|I)$ capture uncertainty about the environment in which the system is to operate. $P(W_j|I), (j = 1 : m)$ define then envelope of weather conditions and $P(F_k|I), (k = 1 : n)$ capture the reliability of the platform and operability aspects related to other aircraft and airspace compliance. Note that we are assuming that W_j and F_k are conditionally independent, namely, $P(W_j, F_k|I) = P(W_j|I)P(F_k|I)$.
- $P(S_i|W_j, F_k, D, I), (i = 1 : l)$ evaluates the quality of autonomous decision making in the GNCC systems of the UAS. This encompasses aspects of robustness and performance of the flight control system, fault detection and diagnosis system, and on-line decisions about reconfiguration of the flight control system and mission rerouting and trajectory planning.

The probabilities $P(S_i|W_j, F_k, D, I)$ are related to the concept of *coverage* discussed in [8], that is, the probability of keeping a desired level of performance given that a particular scenario (weather and fault) has occurred. In the context of this paper, coverage encompasses not only the low-level motion flight controller but, depending on the degree of autonomy of the platform, also the guidance and sense and avoid system.

The probabilities $P(S_i|D,I)$ in (2) are the predicting probabilities of success of the performance index remaining in their region of acceptable performance in one mission. This can be generalised to a cumulative probability of having at least a certain number of successes in a number of missions.

If we would like a single figure of merit for robust autonomy, the natural procedure would be to evaluate the probability that all the indices are jointly within their regions of acceptable performance, namely $P(S_1,\ldots,S_l|D,I)$. This requires the evaluation of the joint conditional probabilities $P(S_1,\ldots,S_l|W_j,F_k,D,I)$. We discuss this aspects further in the companion paper [9].

3 Objective of the Robust Autonomy Certification Framework

The objective of the Robust Autonomy Certification Framework is to provide a mechanism that can allow the NAAs to assess the performance of autonomy without needing to delve into details of its implementation. This is similar to the assessment of a candidate human pilot, which is done based on the candidate's performance in different scenarios rather than neurological aspects. The Robust Autonomy Certification Framework is not attempting to assess if the optimal or best decisions and actions are likely to be made by the implemented autonomy, rather to assess whether decisions made autonomously will result in actions that ensured compliance to pre-established performance measures and expected practices. It is highly likely that there will be many solutions that will ensure compliance and hence ensure an acceptable level of safety.

Decisions made by the NAAs as to whether certify a particular platform for an intended mission depends on two aspects:

- Probabilities of meeting the required performance and safety,
- Consequences of not meeting the required performance and safety.

The framework discussed in the Section 2.2, provides a mechanism for assessing the probabilities. The second aspect of the certification decision depend on the type of mission. A UAS for search and rescue operations at sea may have lesser probability requirements than a UAS for bush fire monitoring operations. If an NAA adopts the proposed framework, then regulations for certification of classes of missions will detailed the required levels or reliability of robust autonomy required.

The evaluation of probabilities will require proprietary GNCC hardware, with its associated software, to be connected to a hardware-in-the-loop (HIL) simulation environment which can simulate different aspects of the mission under various environmental conditions and fault scenarios (sensors, actuators, UAS platform aerodynamics, other aircraft, etc.) This type of evaluation can be under taken by a third party who is independent of any vendor of UAS solutions. This model follows from the current practice in the marine offshore industry used to certify the reliability of the ship positioning and power management systems. For example Marine Cybernetics (www.marinecybernetics.com) is a norwegian company that provides these services for the marine offshore industry, and the evaluation certificates are then presented to the classification societies (the equivalent the NAAs).

Whilst the primary focus of Robust Autonomy Certification is on how the NAAs can certify autonomy, the framework can be also used in the design phase by UAS developers. It would be not only feasible but sound practice to evaluate potential autonomy solutions against the performance required performance measures proscribed by the NAAs. The framework provides insight into weaknesses of particular solutions and allows for integrations that drives the solution towards compliance.

4 Example of Evaluation

In this section, we consider an example, of a test in simulation, of an autopilot part of an RPAS with a capability for control surface fault detection and control reconfiguration [10]. The UAS has a mass of 28Kg and is to be considered for surf-condition monitoring at 1km off the coastline. This aircraft has four control surfaces, ailerons and split elevators, to create redundancy to servo failures via control allocation.

We consider the part of the mission related to approaching a recovery location (before the landing phase). The desired trajectory is shown in Figure 1. The assumed weather conditions are given by a mean wind speed and turbulence spectrum [11], and faults in the four control surface servos are also considered. The environmental conditions for testing and they associates marginal probabilities are summarised in the following:

F_0: Healthy; $P(F_0|I) = 0.7619$.

F_1: Right aileron; $P(F_1|I) = 0.0794$.

F_2: Left aileron; $P(F_2|I) = 0.0794$.

F_3: Right elevator; $P(F_3|I) = 0.0397$.

F_4: Left elevator; $P(F_4|I) = 0.0397$.

W_1: Wind 0 knots; $P(W_1|I) = 0.0909$.

W_2: Wind 10 knots; $P(W_2|I) = 0.6364$.

W_3: Wind 20 knots; $P(W_3|I) = 0.2727$.

To evaluate the performance we consider the indices shown in Table 2. The result of the evaluation of the coverage probabilities $P(S_i|W_j,F_k,D,I)$, is summarised in Table 3. Each probability is computed from data of 100 simulation scenarios - details of these computations are described in [9]. Using these probabilities in (2), we obtain the following measures of robust autonomy:

Table 2 Performance indices and limits for the tested mission.

Index	Description	Limits
r_1	Bank angle	±60deg
r_2	Loading factor	±3.5
r_3	Angle of attack	±11.5deg
r_4	Air speed	$<$ 30m/s
r_5	Horiz Pos Error	±5m
r_6	Vert Pos Error	±5m

- $P(S_1|D,I) = 0.9737$ (Bank angle)
- $P(S_2|D,I) = 0.9737$ (Loading factor)
- $P(S_3|D,I) = 0.9806$ (Angle of attack)
- $P(S_4|D,I) = 0.9820$ (Air speed)
- $P(S_5|D,I) = 0.7600$ (Horiz Pos Error)
- $P(S_6|D,I) = 0.8208$ (Vert Pos Error)

As we can see from these figures, the performance of fault-tolerant autopilot is satisfactory for most performance indices except for the location relative to the desired trajectory, which for this part of the mission has stringent tolerances as specified in Table 2.

To assess what the main limiting issue is, we can analyse the coverage probabilities shown in Table 3. Figure 2 shows a graphical display of these probabilities. These data indicate that the system handles very well the various faults and environmental conditions for the first 4 performance indices (bank angle, loading, factor, AoA, and air speed) with a probability above 0.9 of being inside the region of acceptable performance. The location indices struggle under faults and the degradation of performance is increased with the severity of the weather. Note that for the healthy aircraft the probability of having these indices inside the region of acceptable performance for strong wind conditions drops below 0.8. This suggests that either the controller performance should be improved or perhaps the platform is reaching its operability limits for the desired trajectory. Then performance deteriorates seriously in the presence of faults. This suggests that some improvements can be made in regards of fault detection and handling and or more reliable actuator servos should be considered. Further examination of the simulation scenarios showed that the average time to detect a fault and reconfigure the controller was about 12s. This suggest that improvements in fault-detection could be investigated to reduce the time to detection and system reconfiguration before attempting to replace the servos.

From a point of view of certification, the above six measures of robust autonomy provide a basis information for decision making. The threshold level of acceptance is to be determined by the NAAs. Here the framework only provides probabilities of success, which is only a part of the decision making process for certification. 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 the objective of proposed framework.

Table 3 Coverage Probabilities $P(S_i|W_j, F_k, D)$

S_1	F_0	F_1	F_2	F_3	F_4
W_1	0.9820	0.9410	0.8920	0.9820	0.9820
W_2	0.9820	0.9310	0.9310	0.9820	0.9820
W_3	0.9820	0.9410	0.9410	0.9590	0.9680

S_2	F_0	F_1	F_2	F_3	F_4
W_1	0.9820	0.9410	0.8920	0.9820	0.9820
W_2	0.9820	0.9310	0.9310	0.9820	0.98202
W_3	0.9820	0.9410	0.9410	0.9590	0.9680

S_3	F_0	F_1	F_2	F_3	F_4
W_1	0.9820	0.9500	0.9216	0.9820	0.9820
W_2	0.9820	0.9820	0.9680	0.9820	0.9820
W_3	0.9820	0.9820	0.9820	0.9820	0.9820

S_4	F_0	F_1	F_2	F_3	F_4
W_1	0.9820	0.9820	0.9820	0.9820	0.9820
W_2	0.9820	0.9820	0.9820	0.9820	0.9820
W_3	0.9820	0.9820	0.9820	0.9820	0.9820

S_5	F_0	F_1	F_2	F_3	F_4
W_1	0.9820	0.9750	0.9750	0.9820	0.9820
W_2	0.8450	0.8030	0.7940	0.7050	0.7840
W_3	0.6820	0.0180	0.0180	0.0180	0.0180

S_6	F_0	F_1	F_2	F_3	F_4
W_1	0.9820	0.9750	0.9750	0.9820	0.9820
W_2	0.8450	0.8030	0.7940	0.7050	0.7840
W_3	0.7820	0.7180	0.6180	0.6180	0.5180

5 Conclusions

As the number of potential autonomous operations of UAS increases, so does the need for robust performance assessment as a tool for certification. This papers refines a previously proposed framework, which has a significant resemblance to the methods used for assessment of pilots. That is the method uses scenarios over an envelope of environmental conditions (weather and faults) to make a probabilistic assessment of likelihood of keeping prescribed levels of performance and safety, which are deemed adequate to the missions a particular platform is to conduct.

The probabilistic assessment takes into account uncertainty in the weather conditions and fault scenarios under which the UAS autonomous decision making must operate. Data collected from testing, potentially using hardware-in-the-loop simulations provides information about coverage, namely, the probability that the system will keep adequate levels of performance given a particular weather and fault conditions. The coverage probabilities are then used to compute measures of robust autonomy, which are selected according to the requirements of the mission. These measures are probabilities of satisfactory performance given what has been learned through the system evaluation. The evaluation of performance is done without specific knowledge of the implementation of autonomous decision making, for the same reason that pilot assessment is not done in terms of neurological aspects.

The proposed framework provides the probabilities that are the basis necessary for decision making by the NAAs. If an NAA adopts the proposed framework, then regulations for certification of classes of missions should detailed the required levels or reliability of robust autonomy required.

Whilst the primary focus of Robust Autonomy Certification is on how the NAAs can certify autonomy, the framework can be also used in the design phase by UAS developers. It would be not only feasible but sound practice to evaluate potential autonomy solutions against the performance required performance measures proscribed by the NAAs. The evaluation may provides insight into weaknesses of particular solutions and allow for design iterations that drives the solution towards compliance. We have considered an example of evaluation of an autopilot for an RPAS UAS and, based on the the data collected, discussed how suggestions can be made as to what improvements in either algorithms or hardware could be made for a particular platform.

The focus of this paper is on the use of the framework rather than the details of the computations. The latter is discussed in the companion paper [9].

References

[1] Dalamagkidis, K., Valavanis, K., and Piegl, L., *On Integrating Unmanned Aircraft Systems into the National Airspace system*, Intelligent systems, control and automation, science and engineering, Springer, 2009.

[2] Ducard, G. J., *Fault-tolerant Flight Control and Guidance Systems*, Advances in Industrial Control, Springer-Verlag London, 2009.

[3] Edwards, C., Lombaerts, T., and Smaili, H., (Editors), *Fault Tolerant Flight Control*, Vol. 399 of Lecture Notes in Control and Information Sciences, Springer-Verlag Berlin Heidelberg, 1st ed., 2010.

[4] Perez, T., A. Donaire, and B. Williams, "On Evaluation of Robust Autonomy of Uninhabited Vehicles and Systems" In Proceedings of 7th IFAC Symposium on Intelligent Autonomous Vehicles. Sept., Lecce, Italy, 2009.

[5] Perez, T., A. Donaire, P. de Lamberterie, B. Williams "A Framework for Testing Robust Autonomy of UAS During Design and Certification" Proc of AIAA Infotech Aerospace Conference, March, St Louis, Missouri USA, 2011.

[6] Singpurwalla, N.D., *Reliablity and Risk - A Bayesian Perpective*. John Willey & Sons, UK, Wiley Series in Probability and Statistics, 2006.

[7] Rausand, M. and Høyland, A., *System Reliability Theory: Models, Statistical Methods and Applications*, Wiley Interscience, second edition ed., 2004.

[8] Wu, E. "Coverage in Fault-tolerant Control" Automatica (40), p. 537-548, 2004.

[9] Perez, T., B. Williams and P. de Lamberterie "Computational Aspects of Probabilistic Assessment of UAS Robust Autonomy" 28th International Congress of the Aeronautical Sciences ICAS, Brisbane, Australia, 2012.

[10] de Lamberterie, P. and T. Perez and A. Donaire, "A Low-complexity Flight Controller for Unmanned Aircraft Systems with Constrained Control Allocation" In Proceedings of the Australian Control Conference, Melbourne, 2011.

[11] Keegan, W. B. "Terrestrial environment (climatic) criteria handbook for use in aerospace vehicle development." Technical Report NASA-HDBK-1001, Nasa Technical Standards Program, 2000.

5.1 Copyright Statement

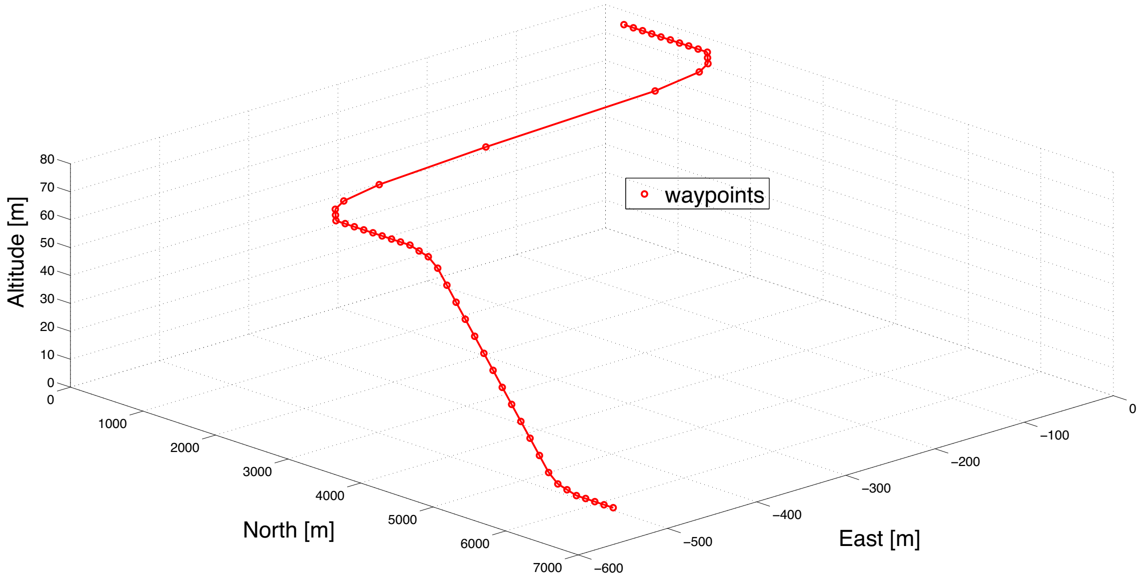

Fig. 1 Desired trajectory for approaching a recovery location.

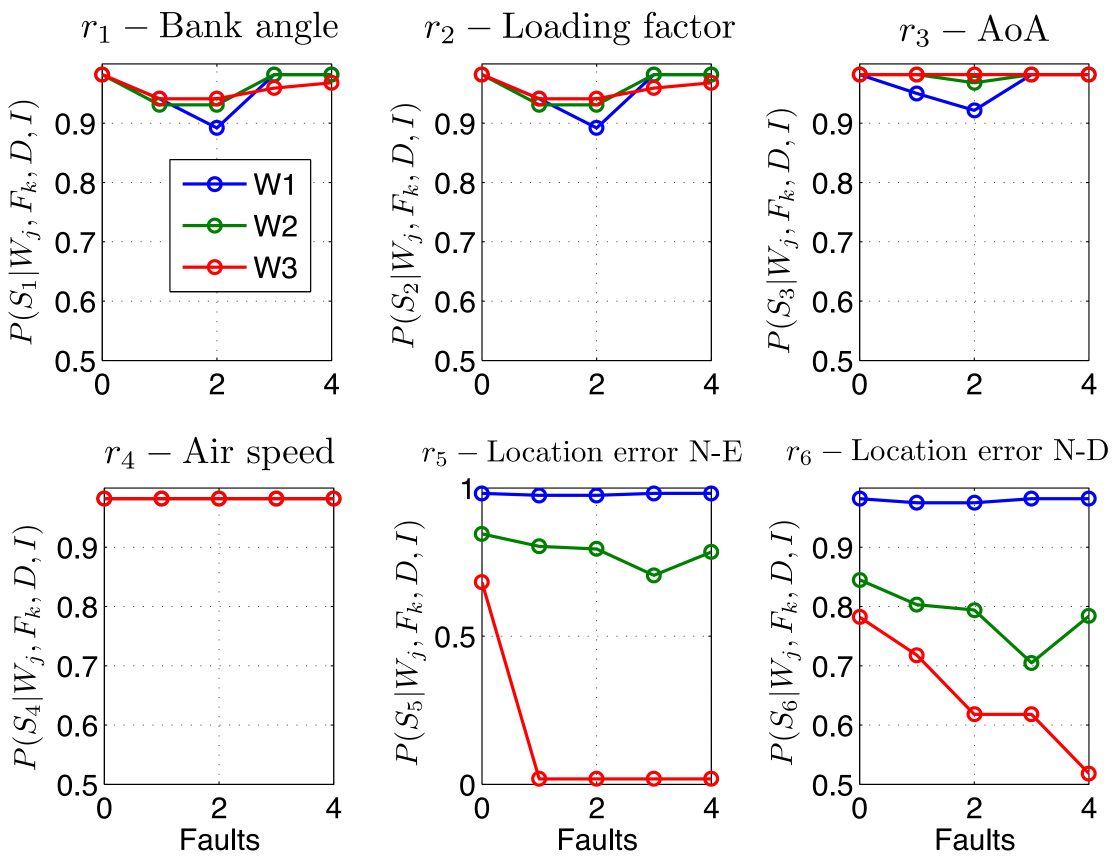

Fig. 2 Coverage probabilities.