**Operational validation of robots for risky environments**

Daniela Doroftei  
Royal Military Academy  
Brussels, Belgium  
Email: daniela.doroftei@rma.ac.be

Eduardo Silva  
INESC TEC  
Porto, Portugal  
Email: eaps@lsa.isep.ipp.pt

Rene Wagemans  
Belgian First Aid and Support Team  
Brussels, Belgium  
Email: rene.wagemans@skynet.be

Anibal Matos  
INESC TEC  
Porto, Portugal  
Email: anibal@fe.up.pt

Vitor Lobo  
Escola Naval  
Lisbon, Portugal  
Email: vlobo@isegi.unl.pt

Geert De Cubber  
Royal Military Academy  
Brussels, Belgium  
Email: geert.de.cubber@rma.ac.be

**Abstract**—This paper presents an operational test and validation approach for the evaluation of the performance of a range of marine, aerial and ground search and rescue robots. The proposed approach seeks to find a compromise between the traditional rigorous standardized approaches and the open-ended robot competitions. Operational scenarios are defined, including a performance assessment of individual robots but also collective operations where heterogeneous robots cooperate together and with manned teams in search and rescue activities. That way, it is possible to perform a more complete validation of the use of robotic tools in challenging real world scenarios.

**I. INTRODUCTION**

One of the problems in the development of outdoor robotic systems (be it aerial systems [1], ground robots [2] or marine platforms [3]) is the lack of adequate test and validation mechanisms to benchmark the performance of the end products. Indeed, it is very hard to quantify this performance in a rigorous scientific manner due to the fact that many variables are out of control in an outdoor environment, e.g. the weather conditions (wind, rain, sea state, illuminance, etc). Moreover, a scientific evaluation requires that multiple trials must be held to validate the statistical significance of the quantitative results, which is not evident when confronted with the evaluation of complex heterogeneous robotic teams in operational conditions, requiring significant logistics for setting up each trial run.

Multiple proposals have been made in the past to remedy this problem. Generally, these validation methodologies can be categorized into two approaches:

A first approach is the development of highly standardized test methodologies [4], e.g. as developed by the National Institute of Standards and Technology (NIST). The advantage of these approaches is that they allow to accurately quantify the robot performance in a number of test setups [5]. However, the disadvantage of these approaches is that, due to their highly standardized nature, these test methodologies are often quite dissociated from practical operational conditions. Figure 1 shows an example of such a standardized test methodology applied on an explosive ordnance disposal robot.

A second approach for validation are robotic competitions (DARPA [7], euRathlon [8], ELROB [9], etc) where multiple robotic systems are pitted against each other in more or less realistic operating conditions. Figure 2 shows an example of such a robot competition (DARPA Grand Challenge) where an autonomous vehicle is navigating through the desert. This validation approach has as an advantage that the performance in real-life like circumstances and environments can be evaluated. However, the disadvantage of these kinds of benchmarking approaches is that, due to their non-standardized nature, they often only allow a qualitative measure of the robot performance and not allow making a detailed quantitative evaluation. An added disadvantage is that coincidence (changing weather and lighting conditions between trial runs, dependence on singular mechanical failures which may not be exemplar for the system operation, etc) plays an important role in these competitions, which may compromise the statistical significance of the benchmarking result.

While both of these approaches are highly valuable and necessary, none of them give an ultimate solution for the
The ICARUS project [10] aims to bridge the gap between the robotic research community and end-users, by developing a toolbox of integrated components for unmanned search and rescue (SAR). The objective of the ICARUS project is to develop robots which have the primary task of gathering data. The unmanned SAR devices are foreseen to be the first explorers of the area, as well as in situ supporters to act as safeguards to human personnel. In order not to increase the cognitive load of the human crisis managers, the unmanned SAR devices will be designed to navigate individually or cooperatively and to follow high-level instructions from the base station [11]. The robots connect wirelessly to the base station and to each other, using a wireless self-organising cognitive network of mobile communication nodes which adapts to the terrain. The unmanned SAR devices are equipped with sensors that detect the presence of humans and will also be equipped with a wide array of other types of sensors. At the base station, the data is processed and combined with geographical information, thus enhancing the situational awareness of the personnel leading the operation with in-situ processed data that can improve decision-making.

Eight different platforms were developed within the project and will be subjected to operational validation tests according to the methodology set out in this paper:

1) A solar airplane which can stay airborne for multiple hours, as shown on Figure 4a.
2) An outdoor rotorcraft which can perform outdoor victim search missions, as shown on Figure 5b.
3) A smaller indoor rotorcraft which can perform indoor victim search missions, as shown on Figure 6a.
4) A large unmanned ground vehicle which can be used to breach through obstacles, as shown on Figure 5a.
5) A small unmanned ground vehicle which can search for victims in smaller voids, as shown on Figure 6.
6) A fast unmanned surface vehicle which can quickly assist victims in the water, as shown on Figure 7a.
7) An unmanned surface vehicle which can provide closer assistance to victims in the water, as shown on Figure 7c.
8) Unmanned rescue capsules which can deploy life rafts to victims in the water, as shown on Figure 8c.

An important aspect within the ICARUS project is that all robotic systems are collaborative agents, which renders the joined evaluation very complex. Moreover, not only the collaboration and interaction between the robots must be validated, but also the joint operation and collaboration with the human search and rescue workers.

III. APPROACH TOWARDS DEFINING THE OPERATIONAL VALIDATION SCENARIOS

As stated above, the ICARUS unmanned tools are meant to be assistive tools for helping the human search and rescue workers to enable them to do their job better, safer and faster. As such, it is of the foremost importance that the robotic tools fit the requirements of the end-users [12] and that the validation methodology is also in line with the real application scenarios as experienced by these end-users. To fill this requirement, a one-week workshop was held where end-users and platform and tool developers were put together in working groups to define a set of use cases for all the tools to be developed. A standardized methodology was followed for use-case redaction [13], leading to a number of use cases. These use cases were then later refined in the form of validation scenarios, taking into account the system requirements and the global definition of the ICARUS demonstration scenarios. During this process, end-users and platform developers were kept in the loop, in order to ensure that the proposed scenarios correspond to realistic platform or tool capabilities and to realistic operational conditions.

IV. OPERATIONAL VALIDATION SCENARIOS

A. Structure

Based on the input of the end users, 10 different validation scenarios were developed. The first scenario, C4I Integration,
is a generic application-agnostic scenario where the integration of the higher-level ICARUS tools in the existing C4I equipment and procedures of search and rescue workers is validated. As no physical robots are involved in this scenario, this scenario is not further discussed in this paper.

All scenarios are chronologically ordered, as depicted on Figure 3 and form, when played one after another, a consistent timeline in line with the demonstration scenarios. Hereby, the leftmost scenario timeline on Figure 3 corresponds to the urban search and rescue (USAR) demonstration scenario, whereas the rightmost scenario timeline on Figure 3 corresponds to the marine search and rescue (MSAR) demonstration scenario. Each of these operational validation scenarios will now be briefly introduced.

B. C4I Mission Planning

In this scenario, the mission planner assigns sectors and tasks to SAR teams. He does this by fusing information from different data sources. This data consists of GIS maps, but also of data from the endurance aeroplane which is tasked to map an area.

Important abilities to be validated by this scenario are:

- The ability to share data with all relevant stakeholders
- Ability for the endurance UAS to deploy quickly and take-off, fly and land safely in difficult weather conditions and on uneven terrain
- Ability to quickly map an area with an UAS
- Ability to quickly overlay acquired geo-referenced visual, infrared and pre-existing map-data
- Ability to assist the mission planner for sectorization and resource allocation

Noteworthy key performance indicators for this validation scenario are:

- Time required for the deployment and flight preparation of the endurance UAS. However, the time required for deployment of and UAS can be impacted by legal constraints of the affected country, which may require specific authorization and/or operator’s qualification in order to operate an UAS.
- Time required from ordering the UAS Area Scan mission (by the mission planner) to visualizing the processed data on the command and control interface.

C. USAR Deployment

In this scenario, the USAR teams move towards and deploy into a sector assigned by the mission planner. The main purpose of this scenario is to test the (rapid) deployment capabilities and the integration of the communication and command and control system. Another purpose of this scenario is to test the network management capabilities when confronted with dynamic team and resource allocations.
Important abilities to be validated by this scenario are:

- Ability to move the unmanned tools without slowing down the team movement
- Ability to quickly deploy the unmanned tools
- Ability for the endurance UAV to detect road blocks and perform a route optimisation mission with the UAS
- Ability to share data with data providers from in the field

Noteworthy key performance indicators for this validation scenario are:

- Difference between the movement speed of the team with and without carrying unmanned tools
- Number of people required to operate the unmanned tools
- Deployment time of the unmanned tools
- Power consumption of the unmanned tools
- Mass, Volume, Battery autonomy, Battery recharge time of the unmanned tools

**D. USAR Apartments**

In this scenario, the USAR team, helped by the large UGV and the outdoor UAS, rescues victims trapped in a semi-demolished apartment building. The main purpose of this scenario is to test the assessment, search and rescue capabilities of the large UGV and the outdoor rotorcraft and their collaborative operation mode.

Important abilities to be validated by this scenario are:

- Ability to search for human victims with the outdoor UAS
- Ability to render (overlay) a map showing geolocalized victim locations
- Ability to assess the medical state of victims
- Ability for the large UGV to remove debris
- Ability for the large UGV to cut through concrete slabs blocking access to the victim
- Ability for the large UGV to place struts to stabilize a structure
- Ability for the rotorcraft to deliver a rescue kit to a victim

Noteworthy key performance indicators for this validation scenario are:

- Victim Search Flight and Processing Time
- Victim Map resolution
- Difference between the real and detected number of victims
- Maximum mass of debris moved by the large UGV
- Time required for breaching through a 2mx2m-sized reinforced concrete slab of 15cm depth
- Time required for outdoor medical assessment

**E. USAR School**

In this scenario, the USAR team, helped by the UGV and UAV systems, rescues victims trapped in a semi-demolished school building. The main purpose of this scenario is to test the assessment, search and rescue capabilities of the small UGV and the indoor rotorcraft and their collaborative operation mode.

Important abilities to be validated by this scenario are:

- Ability for the small UGV and indoor UAS to support various levels of communication bandwidth and various levels of autonomy
- Ability for the small UGV and indoor UAS to perform an indoor structural assessment
- Ability to search for human victims with the indoor UAS and the small UGV
- Ability for the small UGV to set up a communication with victims
Fig. 6: Operational Validation scenarios: School and Warehouse

- Ability to create 2D and 3D maps of combined small UGV and indoor UAS data
- Ability for the small UGV camera to detect fires
- Ability for the small UGV to transport oxygen cylinders
- Ability for the small UGV to find the best exit strategy through explored buildings

Noteworthy key performance indicators for this validation scenario are:

- Standard deviation of the combined 3D map from the ground truth 3D map
- Mean difference between the reported victim position by the small UGV and indoor UAS and the real victim position
- Difference between the real and detected number of victims
- Time required for Victim Search

F. USAR Warehouse

In this scenario, the USAR team, helped by the UGV and UAV systems, rescues victims trapped in a semi-demolished warehouse building. The main purpose of this scenario is to test the assessment, search and rescue capabilities of the small and large UGV and the indoor and outdoor rotorcraft and their collaborative operation mode.

Important abilities to be validated by this scenario are:

- Ability to reconfigure team composition and allocation of unmanned tools
- Ability to control the large UGV manipulator with an exoskeleton
- Ability for the large UGV to breach through walls
- Ability for the large UGV to deploy the small UGV at a height of 1.5m
- Ability for the small UGV to perform an indoor structural and CBRN assessment and victim search operation
- Ability for the large UGV to act as a wireless repeater
- Ability to use the outdoor rotorcraft and small UGV 3D maps to assess the structural integrity of the building

Noteworthy key performance indicators for this validation scenario are:

- Movement speed of the large UGV in GPS waypoint mode through difficult terrain
- Time required for breaching through a 2m x 2m reinforced concrete wall of 15cm depth (min)
- Mean difference between the reported victim position by the outdoor rotorcraft and the small UGV and the real victim position (cm)
- Difference between the real and detected number of victims by the outdoor rotorcraft and the small UGV
- Time required for Victim Search by the outdoor rotorcraft and the small UGV

G. MSAR Air-Air

In this scenario, the MSAR team assesses the situation helped by the endurance aeroplane and deploys into a sector assigned by the mission planner. The main purpose of this scenario is to test the (rapid) deployment capabilities and the integration of the communication and command and control system and the collaborative victim search capabilities of UAS (outdoor rotorcraft and endurance aeroplane). Another purpose of this scenario is to test the network and command and control management capabilities when confronted with dynamic team and resource allocations.

Important abilities to be validated by this scenario are:

- Ability to share mission plans and data with the relevant stakeholders
- Ability to search for and detect victims with the outdoor UAS
- Ability to create a map overlay with geo-referenced visual, infrared and pre-existing map-data
- Ability to assist the mission planner for sectorization and resource allocation
Fig. 7: Operational Validation scenarios: Marine Search and Rescue

- Ability for the outdoor rotorcraft to assess the medical state of victims
- Ability to support multiple simultaneous unmanned rescue operations
- Ability for the outdoor rotorcraft to deploy rescue kits

Noteworthy key performance indicators for this validation scenario are:

- Deployment time of the unmanned tools
- Number of people required to operate the unmanned tools
- Mass, Volume, Battery autonomy, Battery recharge time of the unmanned tools
- Mean time required from ordering the UAS Victim Search mission (by the mission planner) to visualizing the processed data on the command and control interface
- Difference between the real and detected number of victims by the outdoor rotorcraft and the endurance UAS
- Time required for Victim Search by the outdoor rotorcraft and the endurance UAS

H. MSAR Air-Marine

In this scenario, the outdoor rotorcraft tracks one victim on the water. In real time it provides GPS location of the victim to an unmanned capsule. The unmanned capsule autonomously moves towards the victim and inflates the life raft when reaching the victim. The main purpose of this scenario is to test the collaborative victim rescue abilities of the outdoor rotorcraft and the unmanned capsules.

Important abilities to be validated by this scenario are:

- Ability to remote control the unmanned capsule
- Ability for the unmanned capsule to function autonomously
- Ability to provide assistance to 4 victims in the water
- Ability of the command and control system to overlay geo-referenced visual, infrared and pre-existing map-data
- Ability to search for human victims with the outdoor rotorcraft

Noteworthy key performance indicators for this validation scenario are:

- Percentage of victims in the water rescued by the unmanned capsules
- Mean rescue speed (Distance from the take-off location to the victim position divided by the time required to arrive at the victims position) for the outdoor rotorcraft
- Mean rescue speed (Distance from the take-off location to the victim position divided by the time required to arrive at the victims position) for the unmanned capsule

I. MSAR Marine-Marine

In this scenario, the fast unmanned surface vehicle circles an area with visible and thermal cameras pointing inwards. Victims on the water are detected using the cameras images and two unmanned capsules [14] are deployed and head to GPS positions of two victim clusters. These positions are updated in real time from the data collected by the fast unmanned surface vehicle and fed to the unmanned capsules so that they update their motion. Upon reaching the victims, the unmanned capsules inflate the life rafts. The main purpose of this scenario is to test the collaborative victim rescue abilities of the fast unmanned surface vehicle and the unmanned capsules.

Important abilities to be validated by this scenario are:

- Ability for the fast unmanned surface vehicle to function autonomously
- Ability to provide assistance to 4 victims in the water
- Ability to search for human victims with the fast unmanned surface vehicle
- Ability of the command and control system to overlay geo-referenced visual, infrared and pre-existing map-data
• Ability to deploy unmanned capsules from the fast unmanned surface vehicle
• Ability to support multiple simultaneous unmanned rescue operations

Noteworthy key performance indicators for this validation scenario are:

• Percentage of victims in the water rescued by the unmanned capsules
• Mean rescue speed (Distance from the take-off location to the victim position divided by the time required to arrive at the victims position) for the fast unmanned surface vehicle
• Mean rescue speed (Distance from the take-off location to the victim position divided by the time required to arrive at the victims position) for the unmanned capsule

J. MSAR Air-Marine-Marine

In this scenario, the outdoor rotorcraft tracks one victim on the water. In real time it provides GPS location of the victim to the unmanned surface vehicle. The unmanned surface vehicle autonomously moves towards the victim and at a safe distance of 50 m deploys the unmanned capsule. The outdoor rotorcraft keeps tracking the victim and now the unmanned capsule moves towards him and inflates the life raft when reaching the victim. The main purpose of this scenario is to test the collaborative victim rescue abilities of the outdoor rotorcraft, the unmanned surface vehicle and the unmanned capsules.

Important abilities to be validated by this scenario are:

• Ability for the unmanned capsule to function semi-autonomously
• Ability for the unmanned surface vehicle to function autonomously
• Ability to provide assistance to 1 victim in the water
• Ability of the command and control system to overlay geo-referenced visual, infrared and pre-existing map-data
• Ability to deploy unmanned capsules from the unmanned surface vehicle
• Ability to search for human victims with the outdoor rotorcraft
• Ability to search for human victims with the unmanned surface vehicle
• Ability to support multiple simultaneous unmanned rescue operations

Noteworthy key performance indicators for this validation scenario are:

• Number of people required to operate the unmanned tools
• Percentage of victims in the water rescued by the unmanned capsule
• Mean rescue speed (Distance from the take-off location to the victim position divided by the time required to arrive at the victims position) for the unmanned surface vehicle
• Mean rescue speed (Distance from the take-off location to the victim position divided by the time required to arrive at the victims position) for the unmanned capsule
• Mean rescue speed (Distance from the take-off location to the victim position divided by the time required to arrive at the victims position) for the outdoor rotorcraft

V. CONCLUSION

In this paper, different validation scenarios for collaborative search and rescue robots were described. Each of these validation scenarios contains a detailed scenario. Moreover, each validation scenario contains also a list of capabilities which need to be validated. These capabilities correspond to system requirements for the different tools. Finally, each validation scenario contains a score sheet listing a number of metrics which can be used to quantify the performance of the different tools during operational validation tests. As such, it becomes
possible to validate the degree to which each of these system requirements have been attained.

As can be noted, the approach followed here towards validation scenario design and quantitative benchmarking aims to keep a balance between highly standardized (but less realistic) methodologies and highly realistic (but less repeatable) methodologies. Following this methodology, we aim to provide scenarios and quantifiable validation means which are both scientifically relevant and which also ensures the realistic character of the validation setup.

Within the ICARUS project, the validation scenarios presented here are incorporated in 2 demonstration scenarios which will be simulated in summer 2015. Near Lisbon (Portugal), a shipwreck in coastal waters will be simulated, where the Portuguese Navy will intervene to rescue victims in the water, assisted by ICARUS unmanned aerial and marine platforms. In Marche-en-Famenne (Belgium), an earthquake will be simulated and the Belgian First Aid and Support Team will come to the rescue, helped by the ICARUS unmanned aerial and ground vehicles. During these demonstrations, the validation of the different ICARUS tools will be performed according to the methodology presented in this paper.

ACKNOWLEDGMENT

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement number 285417 (ICARUS).

REFERENCES


