Performance Measures Framework for Unmanned Systems (PerMFUS): Models for Contextual Metrics

Hui-Min Huang, Elena Messina, Adam Jacoff
National Institute of Standards & Technology
100 Bureau Drive MS 8230
Gaithersburg, MD 20899
+1.301.975.-3427
hui-min.huang@nist.gov

Robert Wade
US Army AMRDEC
Software Engineering Directorate
Redstone Arsenal, AL
Robert.L.Wade@us.army.mil

Michael McNair
Science Applications International Corp.
8303 N. Mopac Expressway
Austin, Texas 78759
MICHAEL.K.MCNAIR@saic.com

ABSTRACT

In the development of the Performance Measures Framework for Unmanned Systems (PerMFUS), we have established a multiple-axis performance metrics model for the unmanned systems (UMS). This model characterizes the UMS performance requirements by the missions that are to be carried out, the environments in which the missions are to be performed, and the characteristics of the UMS itself. In other words, we focus on the concept of *contextual* metrics and emphasize that performance evaluation and performance specification is based on context.

Categories and Subject Descriptors

J.2 [physical sciences and engineering] unmanned systems performance

General Terms

Measurement, Performance, Design, Human Factors, Standardization, Verification

Keywords

ALFUS, autonomy, collaboration, communication, contextual autonomy, contextual metrics, energy, environment, goal, human-system interaction, HSI, measure, metrics, mission, mobility, perception, power, robot, performance, sensing, task, terminology, test, unmanned system, UMS

1. INTRODUCTION

The Performance Measures Framework for Unmanned Systems (PerMFUS) concept has been described in earlier documents [1][2]. It aims at providing a general framework that establishes sets of metrics, describes an approach, and provides a set of guidelines to facilitate UMS performance measurement. PerMFUS describes how one can organize and analyze the requirements, establish the metrics sets by both instantiating from the established generic metrics and generating additional program-specific metrics, and devise methods to test and evaluate the UMS. The following features of PerMFUS are described in the earlier publications:

1. A three-axis model (see Figure 1). The concept stems from the Autonomy Levels for Unmanned Systems (ALFUS) Framework [3]. Autonomy can be considered an aspect of the UMS performance.
2. A set of performance areas to be focused on in PerMFUS, namely, mobility/navigation, sensing/perception, energy/power, communication, human-system interaction, end-effector, collaboration/coordination, and payload.
3. A systematic approach on how the UMS's hardware and software characteristics contribute to the UMS performance.
4. An initial set of generic environmental characteristics and an initial set of generic metrics.

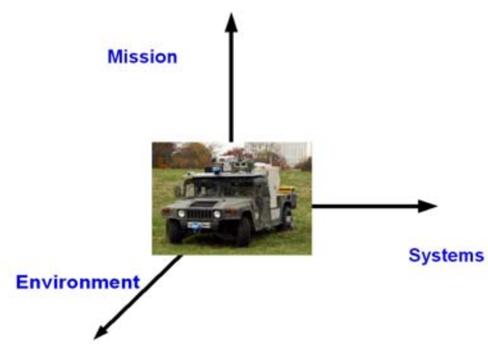


Figure 1: PerMFUS Main Aspects

We continued developing the fundamental features regarding the performance measures of UMSs. In this report, we focus on the concept of contextual metrics. In other words, we maintain that metrics must be associated with certain UMS contexts. For example, speed is a UMS metric, which can be used under different environmental contexts:

- Autonomous speed, teleoperation speed
- Flat/paved surface speed, wet surface speed, speed for climbing a 15-degree hill

We start with describing a set of generic metrics and impose it with various types of contexts.

PerMIS'10, September 28-30, 2010, Baltimore, MD, USA.

2. GENERIC VERSUS SPECIFIC METRICS

A metric is defined as:

> An identified characteristic used to measure a particular attribute of a subject, such as how a defined goal fits a user's needs and whether the system generates required results; a metric can be subjective or objective [4].

We examine the concept of a set of core generic metrics that can be instantiated and applied to different types of task or mission goals. They should also be applicable to most, if not all of the performance areas that PerMFUS describes. The set of metrics includes:

1. Completeness or effectiveness: Is the mission or task goal achieved or to what extent is it accomplished?
2. Accuracy: How close is the mission or task result to the desired or commanded goal state from the perspectives of time, space, and logic; is the result within the desired or commanded tolerance?
3. Efficiency: How much time and/or resources are consumed during the execution of the mission or task?
4. Reliability: What percentage of multiple mission/task executions results in accomplishment of the goals?
5. Safety, integrity, and security: Does the system perform mission/task without the subject being damaged, disrupted, or anyway modified or disclosed un-intentionally?
6. Autonomy: Is the system able to accomplish mission or task goal with minimal human intervention?

These metrics need to be associated with the applicable context to be meaningful. Without the proper context, the UMS's performance data can be either ambiguous or subjected to different interpretations. For example:

- The metric "accuracy" can be applied to mobility and to a temperature sensor, which must be clarified.
- An UMS's specified communication range might be achievable in a line-of-sight situation but not achievable in a non-line-of-sight situation. Therefore, the context of application environment must be stated.

The following sections describe how this common set of generic metrics may apply, at a high level of abstraction, to many types of performance concerns. These generic metrics should be able to be superimposed onto the specific performance metrics for each of the performance areas. They include:

- Navigation/mobility (for unmanned ground vehicles, UGV):
  - Traversing: speed, acceleration, turning radius, brake distance
  - Towing: load size, method
  - Obstacle negotiation: types, severity
  - Stealthiness: signatures of sounds, exhausts, smoke/dusts
- Sensing/perception:
  - Object detection, recognition, location
  - Situation awareness
  - Mapping
- Communication:
  - Range, signal strength
  - Line-of-sight (LOS) versus non light-of-sight (NLOS)
- Energy/power
  - Rates, peak power,
  - Sustained load
  - Endurance
  - Restoration time
- Human-System Interaction (HSI)
  - Controllability
  - Resolutions, update rates of displays
  - Pertaining to Human Supervisor/Operator/Partner
    - Situational Awareness
    - Workload
    - Neglect Tolerance
- End-effector
  - Dexterity, load capacity of manipulator
- Collaboration
  - Information sharing
  - Synchronization

3. MISSION/TASK GOAL-DRIVEN METRICS

Metrics can cover a wide spectrum of issues. However, the focus of the current version of PerMFUS would be on goal accomplishment, as one of the three axes of the PerMFUS model indicates, also highlighted in [2]. An UMS performs missions or tasks. Metrics are required to measure whether and how the mission/task goals are accomplished. The performance of the UMS depends on how the goal is stated. For a navigation or mobility task, the goal can be stated as:

- Go to (x, y)
- Go to (x, y) at time T
- Go to (x, y) after time T
- Go to (x, y) as soon as possible
- Go to (x, y) within (xx, yy) tolerances
- Go to (x, y) by taking the safest route
- Go to (x, y) by taking the shortest route
- Go to (x, y) stealthily
- Go to the nearest covered area

Whether the task goals are completed is measured with different metrics. A quick review finds that the generic metrics can be applied to all these goals.

4. INSTANTIATION TO PERFORMANCE AREAS

Note that the payload performance area is omitted because it can cover too many types of issues.

Performance metrics are required for the various performance areas that PerMFUS identified. Table 1 explores how the generic metrics can be applied.

Table 1: Generic Metrics for Performance Areas

	Mobility	Sensing/ Perception	Comms	Energy/ Power	End-Effector	HSI	Collaboration
Completeness Or Effectiveness	Y/N or %: reached location; covered area	Y/N or %: detected objects; covered area	Y/N or %: transmitted message; covered area	Y/N or %: delivered capacity or peak load;	Y/N or %: reached location; placed objects	Y/N or %: displayed info; enabled control	Y/N or %: on common tasks;
Accuracy	spatial; temporal;	from detection through recognition; mapped covered area	from syntactic through semantic	delivered rated/peak/ sustained energy/ power	location; geometric; temporal; load;	info accuracy; control accuracy (vs. resolution, over/under shoots)	on common tasks;
Efficiency	shorter routes; less obstacles; savings in time, energy, wear and tear, other resources	savings in time, energy, wear and tear, other resources	savings in time, energy, wear and tear, other resources	savings in energy use, wear and tear	savings in time, energy, wear and tear, other resources	amount of displays and devices required; savings in time, energy, wear and tear, other resources	savings in time, energy, wear and tear, other resources
Reliability	% of trials when goal completed	% of trials when goal completed	% of trials when goal completed	% of trials when goal completed	% of trials when goal completed	% of trials when goal completed	% of trials when goal completed
Safety/ Integrity/ Security/	goal accompl-ished without damage or being detected	goal accompl-ished without damage or being detected	info integrity	acquired sufficient energy for tasks	goal accompl-ished without system or object damage or being detected	goal accompl-ished without human or system damage or being detected	goal accompl-ished without damage or being detected
Autonomy	goal accompl-ished with % HSI	sensed or perceived with % HSI	communic-ated with % HSI	energy/ power delivered with % HSI	goal accompl-ished with % HSI	task executed requires % human intervention	goal accompl-ished with % HSI

5. AUTONOMY CONTEXT—MOBILITY PERFORMANCE AREA

Humans have been dealing with various aspects of vehicle performance issues. As such, PerMFUS is set to focus on the unique aspect of the vehicle performance, the "unmannedness." In other words, PerMFUS focuses on a system's performance in the context of autonomy. Meanwhile, it can leverage existing test and evaluation technology that is developed for manned vehicles which do not involve autonomy. For example, there are methods to test and evaluate a vehicle's speed. PerMFUS should focus on only a robot's speed when it is driven without human drivers onboard.

Table 2 should apply to the navigation/mobility performance area.

Table 2: Autonomy Context for Mobility Performance Areas

		Navigation/Mobility Metrics									
		Traversal				Tow		Obstacle Negotiation		Stealthiness	
		sustained speed	acceleration	braking distance	steering radium	load type	tow method	positive obstacle	negative obstacle	noise level	concealment
Autonomy Levels	Remote Control										
	Teleoperation										
	Human-Directed										
	Human-Robot Shared										
	Robot-Directed										
	Fully Auton-omous										

Similar tables can be drawn for other performance areas.

6. ENVIRONMENTAL CONTEXT—MOBILITY PERFORMANCE AREA

The environmental characteristics affect UMSs' performance. Therefore, they should be a part the context of the performance metrics. Table 3 correlates the mobility metrics to the environmental concerns.

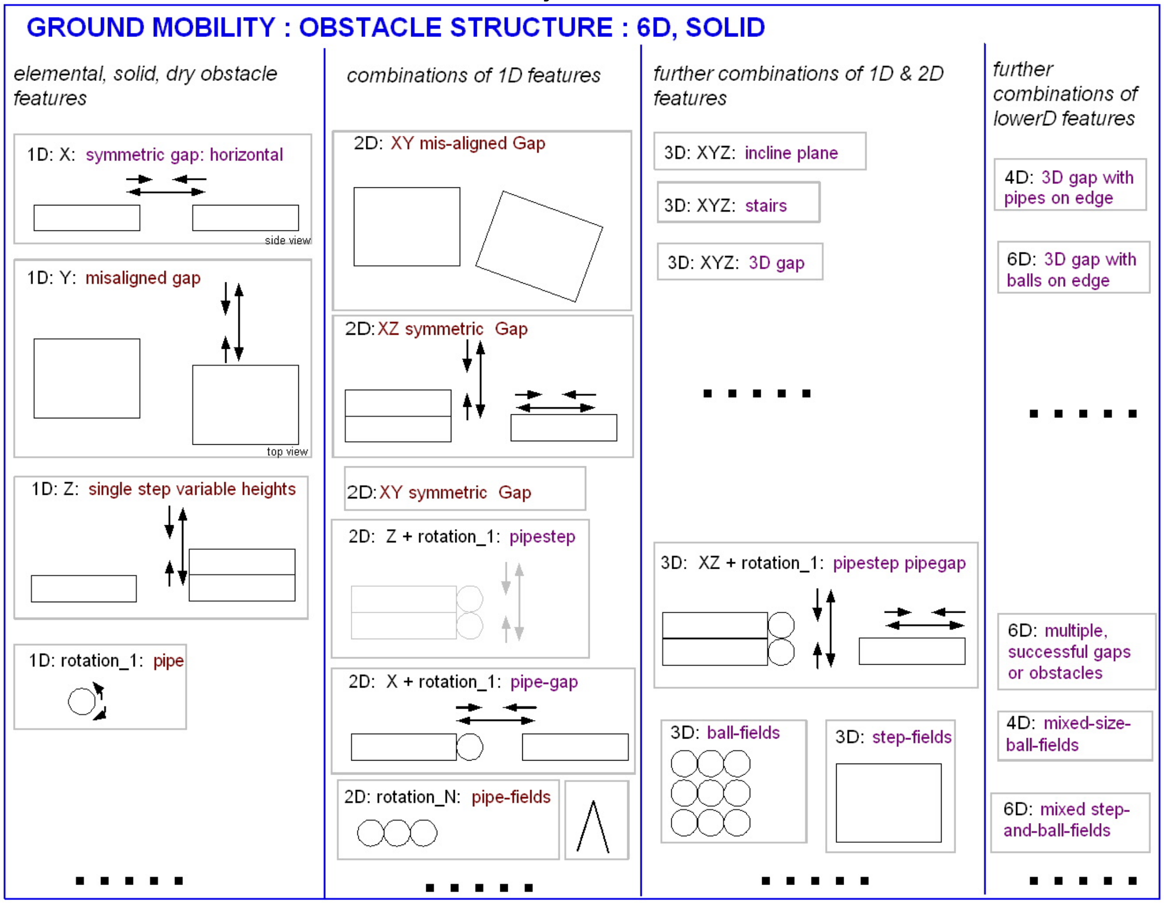
As further exploration, in the environmental classification of positive obstacles for ground UMSs, the following, Table 4 can apply. It illustrates how obstacles can be classified in terms of features such as dimensionality, orientation, complexity, and geometry.

Table 3: Environmental Contexts

		navigation/mobility metrics									
		traversal				tow		obstacle negotiation		stealthiness	
environmental context		sustained speed	acc*	brk dst*	str* rad*	load type	tow method	postv* obs	neg obs	qut*	concl*
aerial	wind										
	lightness										
	rain										
ground	terrain										
	wind										
	lightness										
	rain										
maritime	sea state										
	lightness										
	turbidity										

Key: acc: acceleration; brk: concl: concealment; brake; dst: distance; neg: negative; postv: positive; qut: quietness; rad: radius; str: steering

Table 4: Ground Mobility Positive Obstacle Architecture



7. TOWARD FULL CONTEXT

Besides autonomy and environment, the full context of UMS performance must involve even richer factor descriptions that include missions, tasks, and levels of abstraction. A metric can have different meanings when applied to different levels of abstraction. A laser range sensor can be used to measure a distance at a low level, the same data be used for object recognition at a higher level of abstraction. The issues include:

- Navigation and mobility: A UMS can be commanded in terms of a single explicit position (in some coordinate frame) or sent to an area of concern. Semantic language may be used to describe an area, such as "the other side of this building."
- Sensing and perception: This can range from the pixel level coming out of a sensor up to one through six degrees of freedom; More sophistication can range from detection of entities through object classification; recognition of specific instances of objects ("it's a truck" versus "it's a vehicle") [5].
- Communications: The amount of information transmitted can range from a single data point to composite information. The ability to intelligently plan a communications strategy (e.g., collaborate with other UMSs to form an ad hoc network, save transmissions until out of the tunnel) contributes to the level of autonomy.
- Energy/Power: Management of resource consumption can range from simply reporting current levels to the ability to plan resupplying its own energy or for a team of UMSs.
- End Effector: The manipulation abilities can range from grasping an object to being able to sense and having enough degrees of freedom to allow dexterous assembly of a composite component or handling of delicate, pliable objects.
- Human-Systems Interaction: The level of discourse can range from communicating based on reporting raw data, or from a limited, fixed vocabulary through semantic information. Further up the scale is an UMS's ability to answer questions or formulate specific requests from a human at higher levels of abstraction.
- Collaboration: This can range from simple coordination of mobility to coordinated mission planning and execution among a team of UMSs.

8. CROSS-EFFECTS OF PERFORMANCE AREAS AND UMS SUBSYSTEMS

The ways in which various subsystems might have cross-effects on the areas of performance is an area of study. For example, the mobility subsystem can enhance or impede the coverage areas of sensing and perception as well as the communication. The reverse is true in that the sensing coverage area can affect the mobility and navigation. Further, the degrees of the cross-effects might vary among the various autonomy levels or autonomy modes, making these complex issues.

Table 5 is devised for describing these effects:

Table 5: Subsystem Cross-effects on Performance Areas

		Subsystem Effects On Performance Areas							
		Performance Areas							
		Mobility and Navigation	Communications	Sensing and Perception	Energy/ Power	HSI	End Effector	Collaboration	Payload
UMS subsystem	Mobility and Navigation		traversing for coverage	situation awareness coverage			reach, stability	facilitate	facilitate
	Communications	traverse areas		n/a	endu.	res., usa.	cmd and cntrl	cmd and cntrl	cmd and cntrl
	Sensing and Perception	situation awareness coverage; area reach ability	n/a		endu.	res., usa.	cmd and cntrl	cmd and cntrl	cmd and cntrl
	Energy/ Power	traverse areas; vehicle cntrl	comms avail.	sensing/ percp avail.		avail.	avail.	avail.	avail.
	HSI	cntrllability (C2) for RC and teleop cmding for autonomous	n/a	cmd and cntrl	endu.		cmd and cntrl	cmd and cntrl	cmd and cntrl
	End Effector	n/a	n/a	enhance/ impede sensors	endu.	n/a		cmd and cntrl	cmd and cntrl
	Payload	n/a	n/a	enhance/ impede sensors	endu.	n/a	enhance/ impede reach	cmd and cntrl	

<u>Key:</u>

n/a: not applicable
res: resolution
endu: endurance
cmd: command
cntrl: control
subsys: subsystem
usa: usability
percp: perception
avail: availability

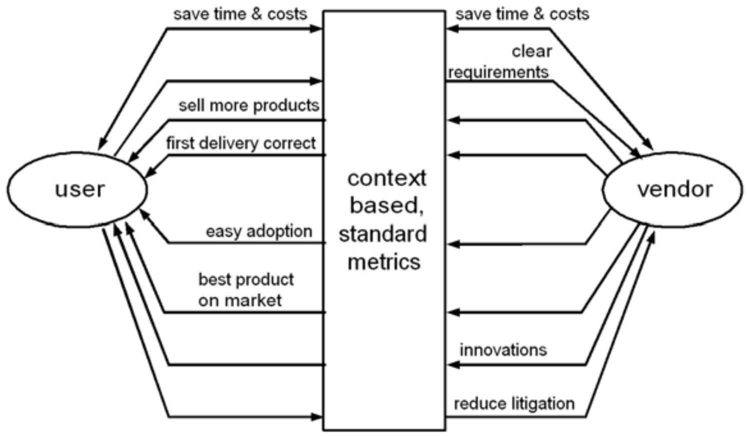


Figure 2: Benefits of Contextual Metrics

9. MULTIPLE ADAPTIVE LIFECYCLES

Besides the testing and evaluation purposes, contextual metrics provide additional benefits to the communication between the users of the UMS and the vendors. These benefits can be illustrated with Figure 2. When the UMS is clearly specified with context and evaluated, the vendors will get clear and unambiguous requirements, they will have a better chance of deliver the UMS right on the first production (thus reducing the costs), the users will have a better chance of acquiring the best-suited UMS on the market as the performance specification is clear to the requirements, and the vendors will have clear technological objectives to provide innovative solutions on their UMS products.

For the relatively new industry of UMS, technologies are evolving and advancing quickly. Some parts of the market lifecycles might exhibit unique features. As users are exploring wider application of the robots, the requirements might evolve as opposed to being established and essentially fixed for the later acceptance testing purposes. In this situation, the following multiple adaptive lifecycles occur:

- The users will be able to evolve and explore advanced requirements to help their operations due to their continuing familiarity with the UMS tools.
- The vendors will be able to devise innovative UMS technology to address the complex requirements.
- The testing and evaluation developers will be able to evolve and enhance the test methods, including the metrics, measures, apparatuses, and procedures, to better address the requirements.
- The users will become more proficient in operating the UMS tools, thus enhancing their mission capabilities.

These lifecycles all iterate with and leverage against each other while advance along their own trajectories.

10. SUMMARY

We described the concept of contextual metrics. Metrics must be associated with proper context to be meaningful. There are many types and layers of contexts that can be associated with metrics. This paper provides a subset. Further development is planned.

11. ACKNOWLEDGMENTS

The authors wish to acknowledge the valuable contributions of the ALFUS working group members and of the participants in the SAE AS-4 Unmanned Systems Committee.

12. REFERENCES

[1] Huang, H., et al., "Performance Measures Framework for Unmanned Systems (PerMFUS)," PerMIS'08 Workshop, Gaithersburg, Maryland, October 2009.

[2] Huang, H., "Ontological Perspectives for Autonomy Performance" PerMIS'08 Workshop, Gaithersburg, Maryland, October 2008.

[3] *Autonomy Levels for Unmanned Systems (ALFUS) Framework, Volume II: Framework Models Version 1.0,* NIST Special Publication 1011-II-1.0, Huang, H. et al., Ed., National Institute of Standards and Technology, Gaithersburg, MD, December 2007.

[4] *Autonomy Levels for Unmanned Systems (ALFUS) Framework, Volume I: Terminology, Version 2.0*, Huang, H. Ed., NIST Special Publication NIST Special Publication 1011-I-2.0, National Institute of Standards and Technology, Gaithersburg, MD, October 2008.

[5] Albus, J.S., et al., 4D/RCS: A Reference Model Architecture For Unmanned Vehicle Systems Version 2.0, http://www.isd.mel.nist.gov/documents/albus/4DRCS_ver2.pdf