Measuring Unmanned Vehicle System Performance: Challenges and opportunities

Harbour, J. L., Bruemmer, D. J., and Few, D. A.

Idaho National Laboratory

(208) 526-4301 (Phone) (208) 526-4311 (Fax)

Jerry.Harbour@inl.gov inl.gov (URL)

Abstract

Technological advances are taking place within the unmanned vehicle system community at an ever accelerating pace. Yet despite such technical advances, there are few published studies that attempt to quantitatively assess the empirical relationship between innovations in unmanned vehicle systems and actual gains in performance. That is, how have (and have not) unmanned vehicle technologies fundamentally improved operational performance. To begin answering this important question, it is imperative that appropriate sets of operationally-defined metrics be developed that quantitatively compare and contrast both "within" unmanned vehicle system performance and "between" manned and unmanned system performance. Ongoing work at the Idaho National Laboratory involving multiple experiments in human-ground robot interaction coupled with "measured" field trials associated with the Autonomous (ground) Robotic Countermine System (ARCM) provide valuable insights into how unmanned system performance metrics may be formally developed and applied. It is suggested that the development, standardization, and acceptance of such quantitative measures will permit better and more informed decisions regarding unmanned vehicle system operational applicability, selection, and benefit, as well as provide a means to trend and track ongoing improvement efforts. Additionally, quantitative metrics can assist government

and private sector organizations alike in better assessing their Return on Innovation Investment (ROI2) in unmanned vehicle system technologies.

Introduction

Technological advances are taking place within the global unmanned vehicle system community at an ever accelerating rate. Yet despite such advances, there are few published studies that attempt to quantitatively assess the empirical relationship between innovations in unmanned vehicle system technology and actual gains in performance. That is, how have (and have not) unmanned system technologies fundamentally improved operational performance over time.

We are reminded of the importance of determining such technology-to-performance relationships or perhaps more correctly, the lack thereof, in a General Accounting Office (GAO) report published shortly after the first Gulf War. The report concluded that, *There was no apparent link between the cost of aircraft and munitions, whether high or low, and their performance in Desert Storm* (GAO/NSIAD-97-134).

As the price of unmanned systems continues to escalate, it is imperative that performance-based metrics are developed that can provide a sound basis for objectively evaluating operational effectiveness. Yet it is equally important that realistic performance expectations for such emerging technologies are set: expectations that are based on a thorough understanding of how technological system performance does and does not improve over time (see Harbour and Marble [2005] for further discussion).

The evolution of the world water speed record as described by Harbour and Blackman (2006) serves as an excellent example of how innovations and associated advances in technology and actual gains in performance correlate as a function of time.

The history of the world water speed record is comprised of three distinct technological 'lineages' as depicted in Figure 1. The depicted lineages represent different types of engines used for speed boat propulsion and include:

- A steam engine, propeller-driven lineage;
- An internal combustion engine, propeller-driven lineage; and
- A jet engine, non-propeller driven lineage.

As illustrated in Figure 1, the first official world water speed record was set in 1874 at a speed of 24.61 mph. Power for this initial record-setting run was provided by a steam engine. As further depicted in Figure 1, speeds improved over time within the steam engine technology lineage via a series of incremental improvements in technology. Such improvements culminated in a final recorded speed record of 45.06 mph that was set in 1903.

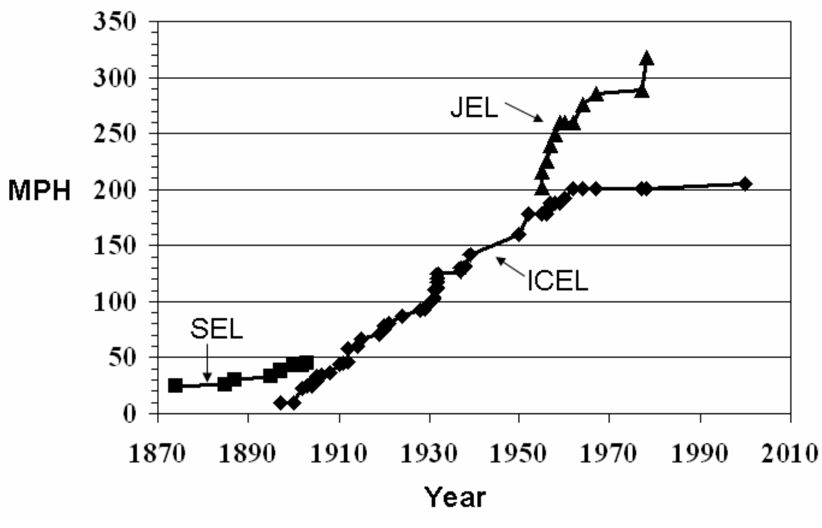

Figure 1. The evolution of the World Water Speed Record is comprised of three distinct technological lineages: the *Steam Engine Lineage* (SEL), the *Internal Combustion Engine Lineage* (ICEL), and the *Jet Engine Lineage* (JEL).

In 1897 an internal combustion engine powered a speedboat that set the first recorded world water speed record for this new technology lineage. The introduction of the internal combustion engine represents a fundamental change or innovation in engine technology. Somewhat surprisingly, however, the addition of this new engine lineage resulted in a net *decrease* in speed of some 29 mph, going from the then current steam engine speed of 39.1 mph to an initial internal engine combustion speed record of only 9.73 mph. One can imagine the jeers and associated condemnation this new 'innovative' technology received when it was first demonstrated.

In fact, it wasn't until 1911, some 14-years later, that the internal combustion engine was able to propel a speed boat faster than a steam engine-propelled craft. With the addition of time and the continual accrual of incremental improvements, however, the internal combustion engine technology lineage eventually increased the speed of watercraft from its modest 9.73 mph beginnings to the current 205.5 mph speed record.

The final innovation to date in the pursuit of the world water speed record is the addition of the jet engine technology lineage, representing yet another innovative change in engine type. The first record for this lineage was set in 1955 at a speed of 202.32 mph. Currently the world speed record is 317.6 mph, recorded in 1978 by Australian Ken Warby while driving the jet-powered *Spirit of America* speedboat.

As illustrated by the world water speed record, remarkable cumulative gains in performance *can* be derived from innovations in technology over time. Yet as noted, such advances commonly proceed by *small* incremental gains as opposed to quantum leaps in performance. Additionally and given enough time, most technology lineages seem to eventually *stall* or plateau out, necessitating the need for the introduction of some new innovation that rekindles and restarts the incremental performance improvement process once again.

Measuring Performance

Performance represents an accomplishment, outcome, or result. Commonly, performance outcome *Y* is expressed as a function of some variable set *X*. A performance metric is simply a measure of how variable set *X* affects performance outcome *Y*.

In the earlier reference to the GAO report concerning the performance of aircraft during Operation Desert Storm, outcome *Y* represented the ratio between fully successful (FS) or 'destroyed' targets and not fully successful (NSF) prosecuted targets. Each plane type (e.g., F-111F, F-117, etc.) was assigned a calculated FS:NFS ratio, allowing comparisons between and among the various aircraft. Based on this utilized methodology, the conclusion was thus reached that aircraft cost and by implication, technological 'sophistication,' was not a critical or key factor in determining performance outcomes as represented by attack effectiveness.

A common challenge in the development of any performance metric system is identifying and developing measures that have both operational meaning as well as practical utility in conducting within- and between-type comparisons. For example we may wish to compare one unmanned ground vehicle (UGV) platform to another, an example of a 'within-type' comparison. Conversely, we may wish to assess how well an unmanned versus a manned system performs the same identical task. This is an example of a 'between-type comparison.

Often both in-process and outcome- or results-based metrics are developed in our attempt to measure and better understand those key factors affecting performance outcomes (Harbour, 1997). Although both types of performance measures are extremely useful, in this paper our focus is only on the development of outcome performance-related metrics as they apply to assessing actual accomplishments of unmanned systems.

Measuring Unmanned System Performance

In reference to unmanned systems, we can create metrics that allow performance assessment at the *component*, *platform*, and *system* levels. For example, we may wish to compare one user control interface to another, an example of measuring performance at the *component level*. In this particular example, we could employ performance metrics of task completion time and navigational accuracy to quantitatively compare and contrast the efficacy of differing interface types.

The work of Nielsen (2006) is used to illustrate this type of component-related performance assessment. Nielsen compared and contrasted teleoperation of UGVs using both 2- and 3-D models of an operator interface. Comparative experimental tasks generally involved finding specific targets in an environment confounded by varying real and simulated conditions. Task speed and accuracy were measured in the conducted experiments. Results indicate that operators generally found more targets faster with the 3-D interface as opposed to the 2-D interface.

We may also wish to compare systems at the *platform level*. For example, we may wish to compare task speed and detection accuracy (e.g., 'hit rate') of UGV platform A to platform B in a landmine detection/marking task using the exact same sensor array.

The use of such speed- and detection accuracy-related metrics further allows us to compare a human doing the same task versus an unmanned system. In selecting specific performance metrics for such tasks it is often important to identify metrics that have practical utility and meaning irrespective of ‘platform’ type (e.g., a robot versus a human). That is, the metrics of task speed and detection ‘hit rate’ are valid performance indicators irrespective if the task is being performed by a robot or a human.

To illustrate measuring performance at the platform level, an Autonomous Robotic Countermine (ARCM) System experiment is described (see Miles and others, 2006 for greater detail). In this experiment, an autonomous-operated UGV attempted to:

- Detect and mark buried anti-tank (AT) mines,
- ‘Report’ or communicate detected mine locations to a distant ground station, and
- Mark a 50 x 1 meter ‘safe’ lane for dismounted troops.

In this example, multiple performance metrics were selected for the experimental trial runs, including,

- Mine detection accuracy (e.g., number of mines correctly detected and expressed as a percentage),
- False mine detection rate (e.g., percentage of false positives)
- Mine marking accuracy (measured in centimeters from the center of the mine),
- Ability to proof and mark a 50 x 1 meter safe lane, and
- Overall task time (e.g., expressed as individual task times per run and as an overall average for all combined runs).

Note that for the ARCM system experimental design, metrics were chosen that allow for both within platform comparisons - that is comparing one unmanned vehicle platform configuration to another - as well as between-platform conditions, thereby allowing quantitative comparisons between human- and unmanned ground vehicle ‘systems.’

Finally we may wish to assess performance at the *system level*, one where multiple unmanned platforms are executing specific missions that require a certain degree of integration and overall coordination. For example in a task that requires physically protecting a wide spatial area against intruder assault, we may wish to assess how the integration of unmanned ground and air vehicles in combination with unattended ground sensors can supplement and enhance a more traditional guns-, guards-, and gates-based protection configuration.

In this system level example, performance metrics are selected that permit comparisons between a predominantly manned-based system *base case* and an unmanned-supported *alternative case*. For this application, it is important to understand how variable X_1, X_2, … X_n, represented by varying types of unmanned vehicle platforms and associated sensor arrays, contribute both individually and collectively to overall system performance under differing operational conditions and configurations.

Specifically in such protection scenarios, it is important to understand how unmanned systems improve (or don't improve) probability of intruder detection, interception, and neutralization, as well as determine how unmanned systems may enhance overall defensive response speed and increase detection standoff distances.

We may further wish to explore the use of unmanned systems to delay intruders during an attack, thus 'buying' greater response time. In this latter example, accrued intruder delay times would also be measured. Only be developing and using such performance measures at the system level can we truly begin to evaluate the potential contributions of unmanned systems in varying protection scenarios.

Summary

Currently in many situations, unmanned air, ground, and water systems have proven their operational effectiveness. In such situations, unmanned systems have successfully accomplished hazardous missions while keeping personnel out of harms way. Yet as the cost and use of unmanned systems continue to rise, it is imperative that we develop performance metrics that can more accurately and quantitatively measure and assess overall unmanned system performance under varying operational conditions. Only by collecting such fact- and performance-based data will decision makers be able to make better decisions regarding unmanned system operational deployment and acquisition. Although it is tempting to simply assume that advances in technology always equate to actual improvements in performance, this assumption may or may not be accurate. Accordingly we must attempt to seek via actual performance measurement that which is actually *true,* as opposed to simply acquiring only what is technologically *new*.

References

Harbour, J. L. (1997). *The Basics of Performance Measurement,* New York, NY: Quality Resources, 71p.

Harbour, J. L. and Blackman, H. S. (2006). Innovation: The other ‘I’ word in performance. *Performance Improvement Journal*, v. 45/2, 24-29.

Harbour, J. L. and Marble, J. L. (2005). How performance improves. *Performance Improvement Journal*, v. 44/8, 14-19.

Nielsen, C. W. (2006). *Using Augmented Virtuality to Improve Human-Robot Interactions,* Ph.D. thesis, Brigham Young University.

Walton, M. C., Bruemmer, D. J., Few, D. A. and Nielsen, C. W. (2006). Developing an Intelligent and Integrated Ground Vehicle System: A case study, Presented at *AUVSI's Unmanned Systems North America 2006 Conference.*