14TH INTERNATIONAL COMMAND AND CONTROL RESEARCH AND TECHNOLOGY SYMPOSIUM

C2 AND AGILITY

EVALUATION OF HIGH RESOLUTION IMAGERY AND ELEVATION DATA

Suggested Topics:
Track 5 – Experimentation and Analysis
Track 8 – C2 Assessment Tools and Metrics
Track 3 – Information Sharing and Collaboration Processes and Behaviors

Walter A. Powell [STUDENT] - GMU*
Kathryn Blackmond Laskey - GMU
Leonard Adelman - GMU
Ryan Johnson [STUDENT] - GMU
Michael Altenau – VIECORE, FSD
Andrew Goldstein – VIECORE, FSD
Daniel Visone - TEC
Kenneth Braswell - TEC

GMU
Center of Excellence in C4I
The Volgenau School of Information Technology and Engineering
George Mason University
4400 University Drive Fairfax, VA 22030-4444 USA
wpowell@gmu.edu
(703) 993-3684

Abstract

How does the underlying data affect the ability of warfighters to derive useful information and make decisions? The Army Topographic Engineering Center (TEC) and GMU endeavor to shed light on this question with the third in TEC's series of value experiments. The ultimate goal of the series is to improve TEC's support of military personnel in the field through improvements made during the development of its geospatial products. The third experiment in the series is fundamentally different from the previous two experiments which presented at the 12th and 13th ICCRTS. In this experiment, instead of evaluating the value of cutting edge geospatial tools while keeping the data constant, we evaluated the effect of higher resolution imagery and elevation data while keeping the tools constant. The high resolution data used was generated from TEC's Buckeye system, an operational airborne surveillance system. This paper discusses the scope of the third experiment, its hypotheses, its experimental design, and our results.

1. Background

As researchers and developers strive to provide advanced tools to process faster and more accurate data, it necessitates the assessment of each innovation so that key resources can be allocated to areas that yield the most "bang for the buck." The JGES program of the U.S. Army Engineer Research and Development Center (ERDC) is designed to meet this need by evaluating the value-added to military decision-making through the use of Geospatial Decision Support Tools (GDSTs) and Advanced Geospatial Decision Support Tools (AGDSTs). GDSTs are computerized tools that display geospatial data and AGDSTs are tools that derive and display based on geospatial data. In theory, GDSTs and AGDSTs can do much more than simply speed up calculations: the potentially superior situation awareness afforded by these tools opens up new possibilities for the conduct of military operations. Translating theory into practice requires a build-test-build cycle that channels technology in spiral development that best support the warfighter. This paper reports on the third in a series of experiments designed to assess the value of Geospatial information to the decision-maker. In the case of military geospatial tools, the ultimate decision maker is the military commander, and the ultimate goal is to support command decisions most effectively.

1. Background

Overhead imagery obtained from aircraft was used extensively for military purposed in World War I and II. Reconnaissance aircraft remained the primary source for overhead imagery until the advent of low earth orbit (LEO) surveillance satellites. Satellites generated non-digital imagery with a 9-25 foot resolution in 1963[1] and had improved the resolution to 6 feet (2 meters) by 1967[2]. Although this resolution was which was significantly less than the 2.5 foot (< 1 meter) resolution available from the U-2 reconnaissance aircraft the area covered by on pass of satellite imagery dwarfed the total imagery collected by surveillance aircraft[3]. The development of digital image technology allowed the military, and later the public, access to digital imagery in the 1-meter resolution range e.g. Google Earth™. While 1-meter resolution imagery was adequate for surveillance and military planning for conflicts involving large equipment and large bodies of men, asymmetric warfare requires the ability to recognize objects with dimensions of less than one meter. The buckeye system was initially developed to provide higher resolution imagery (< 1 meter) than was currently available to facilitate the asymmetric battle through the detection of potential Improvised Explosive Devices (IEDs)[4].

Until relatively recently, elevation data has historical been generated by manual survey. The completion of the space shuttle radar survey generated Digital Terrain Elevation Data level 2 (DTED2) for the earth between 60N latitude to 57S latitude[5]. DTED2 data has an accuracy of +/- 30 meters with data points every 30 meters[6]. DTED2 data is "bare earth" data; it has been processed to eliminate elevation data due to man-made structures (buildings, bridges, other structures) and flora (trees and ground cover). DTED2 data is what is included in most paper and digital topographic maps.

The buckeye system consists of a high resolution digital camera and a Light Detection and Ranging (LIDAR) system to generated elevation data. Buckeye can be mounted on a helicopter or an Unmanned Aerial Vehicle (UAV). Buckeye provides imagery that has 4-inch resolution, is in color, is digital, and is orthorectified (synchronized with known geospatial reference systems). This is imagery is of higher resolution than the best unclassified aerial reconnaissance imagery. The elevation data generated by Buckeye is Digital Terrain Elevation Data level 5 (DTED5) with an accuracy of +/- 1 meter @ 1 meter spacing. Buckeye data is not "bare earth" and accurately depicts the elevation data associated with man-made structures. Buckeye data has been collected for most of the urban areas and major transport arteries in Iraq. Buckeye data is unclassified due to its unclassified source, but is treated as For Official Use Only (FOUO) and is available on the military SIPRNET from TEC.

The goal of the current experiment is to assess the benefits of Buckeye/DTED5 data to the warfighter. Specifically we want to determine the effects of higher resolution data on military decision-making. We will investigate two aspects of how experience military personnel learn from imagery and elevation data: (1) the derivation of information from data and (2) the evaluation of the data. The derivation of information from the available data is learning at the fourth level of Bloom's revised hierarchy of cognitive processes: *xxx* and evaluation is the fifth level: evaluation[7]. The evaluation of imagery and elevation data in a mission specific context requires a series of decisions and judgments based on written policy and experience. This experiment will capture the impact of higher resolution data on the decision-making in a military planning context through quantifying the participant's ability to derive information from and evaluate data.

The paper is organized as follows. Section 2 describes the overall scope of this experiment. Section 3 discusses the primary and secondary hypotheses to be examined. Section 4 lays out the design of the experiment and the reasoning that led to this design. Section 5 discusses the computing environment to be used in the experiment. Section 6 describes the metrics to be used to quantify the results of each trial. Section 7 discusses the results and the impact of the experiment.

2. Scope of Experiment

Our ultimate objective is to evaluate the benefit, to the warfighter, of integrating higher resolution imagery and elevation data with currently available Command and Control planning tools. This third experiment in the program is fundamentally different than the first two experiments, presented at 12th, 13th and this year's ICCRTS conference (Laskey, et.al., 2007, Powell, et.al, 2008, Powell, et.al., 2009). The previous two experiments evaluated the benefits of AGDSTs while keeping the resolution of the data constant while our current experiment evaluates the impact of higher resolution data while keeping the tool set constant.

Discussions with Subject Matter Experts (SMEs) indicated that operational planners for Battalion sized units or larger were unlikely to benefit from Buckeye/DTED5 data. Planners for large units are interested in forests, roads, urban areas, rivers and bridges whereas planners for small units are interested in trees and shrubs, alleys and paths, buildings and walls, streams and fords. Small unit planners would probably benefit more from the Buckeye/DTED5 data than large unit planners. In this experiment the general scenario will ask experienced military operators, working individually and acting as small unit planners, to evaluate multiple potential sites for a Vehicle Control Point (VCP). Follow on experiments will further address different kinds of planning problems, at various levels of command, and involving collaboration among members of staffs as well as individual decision makers.

3. Hypotheses

As we discovered while planning the first experiment, in order to evaluate the military value of AGDST, we needed a clear definition of military value, along with quantifiable metrics of value. Our determination of what constitutes value in this experiment is based on discussions with several experienced military subject matter experts (SMEs). These planners believe that the value of GDSTs lie in their ability to:

(1) Reduce the time spent generating evaluating an area. Since the higher resolution Buckeye data should allow participants to derive less uncertain information, the participants should be able to spend less time subjectively estimating the impact of the uncertainty on the mission. Less uncertain information should also allow the participants to more rapidly form their overall evaluations.

(2) Improve the operator's ability to extract meaningful information. As the resolution of imagery and elevation increases the data approaches reality, and the associated uncertainty decreases. Because the quaintly of the data increased by orders of magnitude, AGDSTs must be used to extract meaningful information. The use of these AGDSTs should allow the operators to extract more meaningful (less uncertain) information.

(3) Improve the operator's ability to generate higher quality mission specific evaluations.

(4) Improve the consistency of planner's evaluations. As the resolution of the available data becomes better, the data becomes a closer representation of truth and information derived from the data becomes less uncertain. Since the participants' judgments are based on better information their evaluations should converge to the most accurate evaluation of each site. This converging should result in more consistent evaluations.

It follows from the criteria above that, in comparison with decision-makers using higher resolution data, we hypothesize that trained and experienced military planners who use Buckeye would:

1. Evaluate the data more quickly. Rationale: Higher resolution data will reduce the uncertainty associated with the information upon which the participant's evaluations are based and thus their evaluations should require less time.
2. Produce higher quality evaluations. Rationale: The higher resolution data should allow the participants to derive better information which should cause their evaluations to agree more closely with the Consensus evaluation.
3. Require less additional information to complete their mission. Rationale: The higher resolution data should provide more information for the participant's evaluation and thus they should require less additional information from external sources.
4. Be able to derive more accurate information. Rationale: When using the higher resolution Buckeye/DTED5 elevation data, the participants should be better able to use the tools inherent in CSE to derive answers to questions about the terrain.
5. The evaluations based upon the Buckeye/DTED5 data would be more uniform i.e. have less variance in two of the four categories above (better

information and better evaluation) than those evaluations based on CIB1/DTED2 data. Rationale: Using higher quality information derived from less uncertain data should cause the participants evaluations (each criterion and overall) and to agree more closely.

As the determination of military value and the design of the experiment evolved, we identified two secondary hypotheses. First, the structure of the experiments requires the repetition of evaluations and there was concern that a learning effect my might skew the results of the experiment. Second, although not a concrete benefit the perception of the participant as to the specific utility of the Buckeye data will assist in the integration of Buckeye data into deployable systems. The secondary hypotheses investigated include:

6. There would not be a learning effect due to experimental design. Rationale: The participants have previous training and experience using C2 planning tools and the tasks the participants are asked to perform are similar to those that they have performed in the normal course of their duties.
7. Participants will consider using Buckeye superior with respect to speed, ease of use, usefulness of information, and overall. Rationale: The participants should consider the high resolution Buckeye's data of benefit in the planning process.

4. Study Design

The study design employs a factorial design with four independent variables: Data Source (Buckeye or CIB1), Data Order (whether the first scenario is worked with Buckeye data or CIB1 data), Scenario Order (whether scenario 1 or 2 is worked first) and Data Type (imagery or elevation data). Data Source and Data Type will be within-subjects variables because each participant will work one independent planning scenario with CIB1/DTED2 and one independent scenario Buckeye/DTED5 data and each participant will use both imagery and elevation data. A within-subjects design is particularly valuable when the number of available participants is limited, as in the current case. Results from the sets of tasks can be compared for each participant, thus eliminating participant-specific effects that might add variability to the results. Data Order and Scenario Order will be between-subjects variables because any given participant can only experience one ordered sequence for these variables without repeated exposure to both data sources.

The participants will evaluate potential VCP sites using the same underlying C2 system, the Commander's Support Environment (CSE), but using two different sources for the imagery and elevation data. One of the evaluations will be performed with Buckeye/DTED5 data and the other evaluation will be performed with CIB1/DTED2 data. A third evaluation will be performed on the sites which were originally evaluated with CIB1/DTED2 but this time augmented with Buckeye/DTED5 data. All the trials are essentially identical except for the use of source of the imagery and elevation data. The Data Source for the evaluations will be randomly selected so that half of the participants use Buckeye and DTED5 first. Randomizing the order of the tasks will enable the analysis to control for learning effects.

The instructions, sites, evaluation criteria, and tools are the same in both scenarios with the exception of geographic references necessitated by the requirement to have different geographic areas for each trial. Different geographic areas are required to prevent participants from repeating their responses from the first scenario when they form responses for the second scenario. The trios of sites have been carefully selected for their geographic similarity such that the evaluations performed by the participants and the expected results will be as nearly identical as possible. Randomization will be used to control for differences between scenarios.

The participants will be Army enlisted personnel and officers who have previous experience establishing VCPs in Iraq or Afghanistan. They will be split into two groups that are as evenly balanced with respect to ability rank/time in service as possible. Further evaluation of the relative ability/experience of the participants was not possible due to the inability to contact the participants prior to conducting the trials. Group I will perform the evaluations first Buckeye/ DTED5 and then with CIB1/DTED2. Group II will perform the evaluations in reverse order. The groups will be further divided into two subgroups while maintaining the balance of ability and knowledge. Each subgroup will perform the same evaluations for the same two scenarios, but the two subgroups will see the two scenarios in the opposite order. This design will allow us to control for differences due to the order of system use and the scenario order. Each participant will conduct a third evaluation; this evaluation will be on the set of sites on which the participant originally used CIB1/DTED2 data but in this evaluation, the participant will have access to the Buckeye/DTED5 data as well.

The evaluation will consist of evaluating one of two sets of three similar potential VCP sites (a scenario) on the same 28 criteria in six categories. In each evaluation, the participants will evaluate each site on the 28 criteria with respect to (1) the potential of the physical site for establishing a temporary VCP and (2) the amount of additional information that would be required to actually establish a VCP at each site. For each site, the participants will also answer four questions that will only require that they derive information using the tools inherent in CSE. After completing the evaluation of all three sites in a scenario, the participants will rank the sites relative to one another on the overall quality the site for a VCP and estimate their confidence in their ranking. The participants will conduct one evaluation using CIB1/DTED2 data, one using Buckeye/DTED5 data and third evaluation which revisits the site evaluated with CIB1/DTED2 data but with the Buckeye/DTED5 data also available. After completing all three evaluations the participants will weight the relative importance of the categories and criteria and complete a questionnaire comparing the relative benefits of Buckeye/DTED5 and CIB1/DTED2 in the areas of speed, ease of use, utility of the information, and overall.

The evaluation of the participants' decisions will be evaluated verses a "consensus" score from three SMEs *or the global average*. The SMEs, U.S. Marines, a Captain, a Staff Sergeant, and a Corporal, had access to all the available data for each potential VCP site. They evaluated each potential VCP site on all 28 criteria individually and then compared their ratings. After discussing the factors that may have lead to any differences in their evaluations, the SMEs generated an agreed upon "Consensus" score for each criterion for each site. Although the sites were chosen to be similar, the variation due to the differences in the sites needs to me minimized. The absolute difference between the participants' scores and the consensus scores will be used as the as the basis for our statistical analysis. The use of the absolute difference will minimize the variation due to the differences among the sites.

The computers used for the experiments were not homogeneous; four Dell desktops, four Dell XPS laptops, two Prostar laptops, and two other Dell laptops were available. The laptops were configured such that the monitor resolution and area displayed were as near to, but not less than, that of the desktops. All the computers were dual core with greater than 2.0 GHz processors and USB 2.0 capability. The laptops were provided with mice so that participants would no be required to use touchpads. In order to minimize any variation due to the participants using specific computers, the participants use randomly assigned computer for each trail.

Prior to beginning the tasks, both groups of participants will receive standardized training on the use of CSE. The training will be sufficient to allow the participants to perform the required evaluations and will include training on the tools and features unique to CSE. The last phase of the training will require the participants to perform the complete evaluations of two training sites similar to those that the participants will encounter during the trials.

5. Environment

The evaluation will be conducted using the Commanders Support Environment (CSE) as the Command and Control (C2) planning system. CSE is a robust C2 planning and execution based system developed for experimentation. The CSE was originally developed for Defense Advanced Research Projects Agency (DARPA)/Army Multi-Cell and Dismount C2 Program (M&D C2) which was continued from the FCS C2 program. M&D C2 program hosted a series of experiments designed to test out network centric warfare concepts. The CSE is primarily written in C++ code for the Microsoft Windows environment. It is built upon the Viecore FSD Decision Support System (VDSS), and the Data Analysis and Visualization Infrastructure for C4i (Davinci) Toolkit. The VDSS architecture enables the quick addition of modules for communication between CSE and other systems and components. The CSE's GIS components are built upon the Commercial Joint Mapping Toolkit(C/JMTK) which includes ESRI's ArcGIS Desktop licensed at the ArcEditor level.

The CSE provides one main AGDST's, an optimized Line of Sight (LOS) analysis tool, in addition to displaying imagery and elevation data. The LOS AGDST displays a real-time analysis based on the relevant digital elevation data. The display can include a 360^o fan or an elevation cross section out to 5 km from the cursor.

6. Metrics

The participants were all assigned participant numbers and evaluation designators so that their evaluations would be anonymous. The participants' record their evaluations of each criterion on 5-point Likert scales and by using their evaluation designators they will be indistinguishable as to participant. The criteria for evaluation of the Buckeye/DTED5 data will be (1) a comparison of the rapidity with which the evaluation are conducted, (2) the quality of those evaluations, (3) the accuracy of the derived information, and (4) the perception of the participants regarding the merits of the Buckeye/DTED5 data.

Time to Completion. The evaluation of how quickly the participants complete their evaluations will be measured objectively by logging the amount of time it takes participants to complete the tasks. The maximum duration of each trial will be 1.5 hours. The actual time will be calculated by taking the difference between the start and stop times and subtracting any break time.

Evaluation Quality. The evaluation of the participants' evaluations will be objective. We will consider three factors that contribute to the quality of a participant's output: (1) how well the participant's Area Characteristic scores agree with the Consensus *and how well they agree with the global average*; *(2)* how well the participant's RFI scores agree with the Consensus *and how well they agree with the global average*; (3) how well the participants' site ranking agree with the Consensus *and how well the agree with they global average*; and (4) The change in both the Area Characteristic and RFI scores between CIB1/DTED2 evaluations and the revisit evaluations with Buckeye/DTED5. The participants will evaluate each site on 28 criteria in six categories. The scoring will be done on a 5-point Likert scale. The absolute difference between the participants' scores and the Consensus *or Global Averages* will be used to as a measure of agreement with each.

Accuracy of derived information. We asked the participants to answer four questions about each site. Two of the questions require that the information be derived from an examination of the imagery and two require that the participants use the LOS AGDST (acting on the elevation data) to derive the information. Unlike the site evaluations, these questions are objective in that there is a right answer and require little or no subjective analysis on the part of the subjects. Deriving this information is typical of the many individual tasks that are required to make the overall evaluations. The information generated by the subjects will be compared to "ground truth" answers derived using all the available data.

Perception of Merits. We will administer a questionnaire to evaluate the participants' subjective judgment of the benefits of Buckeye/DTED5 data as compared to CIB1/DTED2 data. The participants evaluate which data is more beneficial as to speed, ease, and value of information with respect to the imagery and the elevation for nine tasks and overall. Like the evaluation criteria, the participants will evaluate the participants' answers on a 5-point Likert scale. The results of these questions and a hot wash-up conducted at the conclusion of the experiment are particularly valuable in guiding the future integration of Buckeye/DTED5 data with deployed systems.

7. Analyses of Results

(2005). "“BuckEye” Fact Sheet An Airborne High Resolution Digital Imaging System." Retrieved October 25, 2008, from http://www.tec.army.mil/fact_sheet/Buckeye_factsheet.pdf.

Anderson, L. W., D. R. Krathwohl, P. W. Airasian, K. A. Cruikshank, R. E. Mayer, P. R. Pintrich, J. Raths and M. C. Wittsrock (2001). A Taxonomy for Learning, Teaching, and Assessing. New York, Longman.

Pike, J. (December 3, 2008). "Digital Terrain Elevation Data [DTED]." Retrieved December 3, 2008, from http://www.fas.org/irp/program/core/dted.htm.

Rabus, B., M. Eineder, A. Roth and R. Bamler (2003). "The shuttle radar topography mission--a new class of digital elevation models acquired by spaceborne radar." ISPRS Journal of Photogrammetry and Remote Sensing **57**(4): 241-62.

Richelson, J. T. (2003). Civilians, Spies, and Blue Suits: The Bureaucratic War for Control of Overhead Reconnaissance, 1961-1965, A National Security Archive Monograph,CIA, NRO, DAF.

---

[1] Richelson, J. T. (2003). Civilians, Spies, and Blue Suits: The Bureaucratic War for Control of Overhead Reconnaissance, 1961-1965, A National Security Archive Monograph,CIA, NRO, DAF.

[2] Ibid.

[3] Ibid.

[4] (2005). "“BuckEye” Fact Sheet An Airborne High Resolution Digital Imaging System."   Retrieved October 25, 2008, from http://www.tec.army.mil/fact_sheet/Buckeye_factsheet.pdf.

[5] Rabus, B., M. Eineder, A. Roth and R. Bamler (2003). "The shuttle radar topography mission--a new class of digital elevation models acquired by spaceborne radar." ISPRS Journal of Photogrammetry and Remote Sensing **57**(4): 241-62.

[6] Pike, J. (December 3, 2008). "Digital Terrain Elevation Data [DTED]."   Retrieved December 3, 2008, from http://www.fas.org/irp/program/core/dted.htm.

[7] Anderson, L. W., D. R. Krathwohl, P. W. Airasian, K. A. Cruikshank, R. E. Mayer, P. R. Pintrich, J. Raths and M. C. Wittsrock (2001). A Taxonomy for Learning, Teaching, and Assessing. New York, Longman.