Scythe

Proceedings and Bulletin of the International Data Farming Community

Issue 8 - Workshop 20

Proceedings and Bulletin of the
International Data Farming Community

Table of Contents

Scythe

Proceedings and Bulletin of the International Data Farming Community

It is appropriate that the publication supporting the International Data Farming Workshop is named after a farming implement. In farming, a scythe is used to clear and harvest. We hope that the "Scythe" will perform a similar role for our ***data*** farming community by being a tool to help prepare for our data farming efforts and harvest the results. The Scythe is provided to all attendees of the Workshops. Electronic copies may be obtained from harvest.nps.edu. Please contact the editors for additional paper copies.

The *Scythe* consists primarily of team reports written by the team members reporting on activity in their team during the workshop ***from their perspective***. Please let us know what you think of this eighth prototypical issue. Articles, ideas for articles and material, and any commentary are always appreciated.

Bulletin Editors

Gary Horne: ....gehorne@nps.edu

Ted Meyer: ....temeyer@nps.edu

International Data Farming Community Workshop 20 Program Committee

International Data Farming Community Overview

The International Data Farming Community is a consortium of researchers interested in the study of *Data Farming*, its methodologies, applications, tools, and evolution.

The primary venue for the Community is the biannual International Data Farming Workshops, where researchers participate in team-oriented model development, experimental design, and analysis using high performance computing resources... that is, Data Farming.

Scythe, *Proceedings and Bulletin of the International Data Farming Community,* Issue 8, Workshop 20 Publication date: June 2010

IDFW 20: Enrichment

by Gary Horne
Naval Postgraduate School

We were pleased at the Naval Postgraduate School to once again host an International Data Farming Workshop. Number 20 was held in Monterey from March 22nd through 25th, 2010. This workshop coincided with "Enrichment Week" (hence our theme!) at NPS, a time when students could take a week to explore something different from their normal course of study. We had a record number of teams (seventeen) and our format was a bit different with poster sessions to start us off on Monday morning and outbriefs concluding our work on Thursday afternoon. But our goal during the abbreviated week was as usual, to work in teams using data farming methods to explore our important questions. And on behalf of the SEED Center for Data Farming at NPS, I would like to express our thanks to the team leaders, the plenary speakers and all of the participants in IDFW 20!

This issue, our eighth, of The Scythe contains a summary of each work team effort. And, as always, the plenary session materials, in-briefs, and out-briefs from this workshop are available online at http://harvest.nps.edu along with electronic copies of this issue of The Scythe. The plan continues to be to hold even-numbered workshops once a year in Monterey with odd-numbered workshops taking place at international venues. So looking ahead, our Data Farming community will be in Lisbon, Portugal for our next workshop, International Data Farming Workshop 21. I would like to invite you to participate, starting with the pre-workshop dinner on September 19th, 2010. The workshop will be held from September 20th through 24th and we hope to see you there!

Gary Horne

International Data Farming Workshop 20
March 2010 — Monterey, California

Team 1: The Logistical Impact of Marine Corps Enhanced Company Operations

TEAM 1 MEMBERS

Capt Daniel Hinkson, USMC
Capt Chad Puff, USMC
Mary McDonald
Naval Postgraduate School, US

Richard McCourt
Defence Research and Development Canada

Capt Kyungtack Oh, ROK Army
University of Texas at Austin, US

INTRODUCTION

The United States Marine Corps uses maneuver warfare as a basic doctrinal concept to fight its battles. Maneuver warfare demands the ability to avoid the enemy's strengths and attack his weaknesses in ways that are advantageous to the overall strategy. This overarching concept has heavily influenced the development of the Marine Air Ground Task Force (MAGTF)—a size-scalable, combined-arms, multi-mission-capable force used across the spectrum of conflict. The Marine Corps is continually developing tactics, techniques, procedures, and technologies that seek to increase the efficiency and lethality of the MAGTF. In this spirit, the Marine Corps Warfighting Laboratory (MCWL), whose "purpose is to improve current and future naval expeditionary warfare capabilities" (MCWL Website), is currently exploring the viability of a concept called Enhanced Company Operations (ECO).

In August 2008, the Commandant of the Marine Corps, General James Conway, signed a white paper entitled, "A Concept for Enhanced Company Operations," which states:

Enhanced Company Operations describes an approach to the operational art that maximizes the tactical flexibility offered by true decentralized mission accomplishment, consistent with commander's intent and facilitated by improved command and control, intelligence, logistics, and fires capabilities. Enhanced Company Operations will be reliant on increased access to, and organic control of, functional support, as well as excellence at the individual, squad, and platoon levels. As such, it builds on the results of Distributed Operations experimentation and capability development to provide battalion commanders the critical link between operational planning and squad level tactical execution.

ECO involves reorganizing and augmenting the traditional rifle company in a manner that contributes to "enhanced" C2, intelligence, logistics, and fires capabilities. This process not only involves personnel changes, but also specific training and technological improvements. The end state is to develop the company's ability to become the base maneuver element of the MAGTF, a role traditionally held by the infantry battalion. Changes include the incorporation of a company-level operations center, a company-level intelligence capability, enhanced fire support coordination, and personnel specifically tasked to focus on logistics.

This team participated in an ongoing Naval Postgraduate School thesis project to explore the logistical impact of a deployed enhanced company on a Marine Expeditionary Unit's (MEU) supporting assets. At the start of the workshop, the team had the following goals:

1. Assess and refine a simulation model developed using Map-Aware Non-uniform Automata (MANA) to evaluate the logistical impact of Marine Corps Enhanced Company Operations on a Marine Expeditionary Unit.
2. Determine appropriate variables and ranges to incorporate into an experimental design.

The MANA model referred to above is based on a realistic Africa-based scenario that allows enemy agents to influence the logistical demand of the supported company.

Description of Scenario

This study uses a scenario developed by MCWL and used during the ECO Fires Conference of 21-23 April 2009, which provides a realistic operational environment with which to test the ECO concept. The fictional scenario takes place on the African continent in the border area between Burundi and the Democratic Republic of the Congo. In the notional orders describing the scenario, MCWL changed the names of the countries to prevent others from mistaking them for real-world events.

The United States has a supportive relationship with the government of Bunduri, a relatively stable democracy in East Africa. The U.S. has a neutral relationship with the government of Razie, which is led by a corrupt president who has used various nefarious means to stay in power for many years. Within Razie, there is a government opposition movement called the Movement for Democratic Change (MDC). In the latest elections, the leader of the MDC won the popular vote, but the sitting president refused to recognize the election results. As a result of internal and international pressures, the two parties reached a power-sharing agreement with the president remaining in place and the winner of the elections serving as prime minister.

After a failed assassination attempt on the prime minister, in which the president's followers were implicated, the president dissolved the national government and instituted martial law. The prime minister fled to the east of Razie with his MDC followers. The MDC's military arm, the Manicaland Peoples Force (MPF), rebelled and took control over Manica Province in Eastern Razie. The former prime minister announced the formation of the independent state of Manicaland and declared war against Razie. Additionally, he declared Manica tribal lands within Bunduri as a part of

Manicaland, and the MPF crossed into western Bunduri. The contested area is shown in Figure 1.

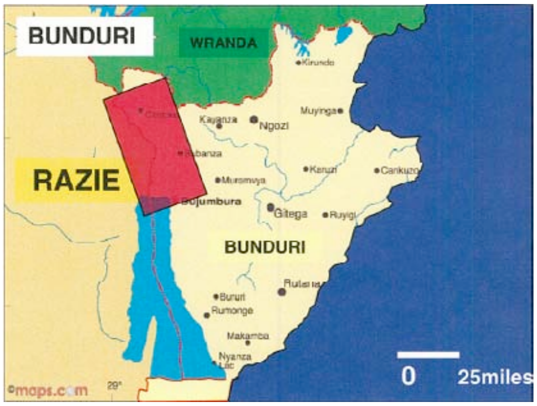


Figure 1: Scenario Area of Operations

The simulation attempts to model ECO-capable Alpha Company. MPF forces have been kicked out of Alpha Company's area of operations (AO), but they continue to make incursions across the border to influence the local populace and to harass friendly forces. Since Alpha Company is the main effort, they have the luxury of receiving the priority of support from the MEU's assets.

The MANA Model

This team began the conference with an initial model representing Alpha company's AO already built in MANA version 5. A screen shot is shown in Figure 2.

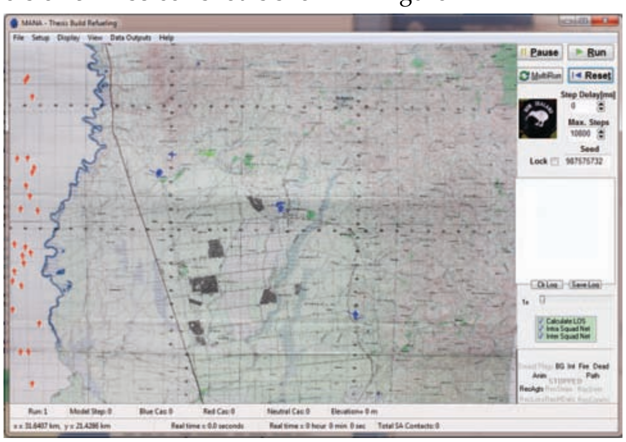


Figure 2: Screen shot of initial MANA model

The model background is a topographical map of the scenario area. The red agents start on the Razie side of the boarder (denoted by the blue river) and will attempt to make it to the right most side of the play board. Alpha Company has three established platoon positions within their AO. At each position, two squads run patrols, and one squad remains for security. There is also a 60mm mortar team at each platoon position. As the model runs, the patrolling blue agents will interdict the red agents when they come into sensor and weapons ranges.

WORKSHOP RESULTS

The team spent the first day of the workshop familiarizing themselves with ECO, the scenario, and MANA.

The team spent the second day attempting to incorporate the scenario logistical aspects into MANA. Our original idea was to have a supply "tank" at each platoon position that would hold two days of supply units for each blue agent at that position. Each agent starts with one day of supply, and returns to the patrol base for a resupply when its tank is empty, so each agent has a total of three days of supply available at the start of the scenario. When the patrol base tank is depleted, a resupply agent, the MEU helicopter, flies to the tank and replenishes the supply units. We attempted to model this behavior using MANA's fuel parameter and auto-refueller agents.

Our initial attempt at implementing this scenario had the agents transition from their default state to a "fuel out" state when their resources were depleted, and then to a "refuel by friend" state when they come into contact with the resupply tank. But this process did not give the desired behavior, because as each squad entered the "fuel out" state, some of the agents would immediately transition into the "refuel by friend" state due to their close proximity to other friendly agents in the squad. Some agents transferred back to their default state without receiving any fuel, which then continued to decrement their fuel parameter below zero, causing the agent to never enter the "fuel out" state again. Many different combinations of triggers and trigger states were attempted in order to get this refueling scenario to work, but no combination produced the desired result.

The team spent day three trying to create the desired logistical behavior. This time, instead of using MANA's refueling states, we relied on the different sensor and weapon parameters to trigger when an agent could refuel itself by using a negative fuel consumption rate. For instance, an agent starts patrolling with one day of supply units. Once those units are depleted, that agent changes to a state that is visible to the resupply tank's sensor and returns to the base. When the agent returns to the base, and comes within the resupply tank's weapon's range, it is shot by the tank. The agent switches into the "shot at" state, refuels itself, and then continues on its mission. When the tank has fired all of its ammunition, which is used to represent the tank's supply capacity, it changes states into one visible by the resupply agent (the MEU helicopter), is shot at by the resupply agent, and reverts back into the default state with a full load of ammunition (i.e. supplies).

While this algorithm worked in a simplified model,the desired result could not be achieved with a more complicated scenario. Our conclusion was that model limitations in the version of MANA used at the workshop made it unsuitable for this particular study. But after the workshop, Capt Hinkson contacted one of the MANA developers, Mark Anderson, who explained that MANA uses a random draw to determine which agent gets to have its turn first. MANA also did not implement trigger state changes in the same time step in which they occurred, which explained the behavior we observed. Mr. Anderson provided Capt Hinkson with an updated version of MANA 5 that included instantaneous state changes. The model now appears to be working as desired and this version was used in further work to complete Capt Hinkson's thesis research.

Team 2: Using Data Farming to Examine Various Aspects of the Transformable Craft

TEAM 2 MEMBERS

Dr. Santiago Balestrini-Robinson
Georgia Institute of Technology

Mr. Paul T. Beery
Naval Postgraduate School, US

Maj. Huntley Bodden
Naval Postgraduate School, US

LT Robert Cizek
Naval Postgraduate School, US

LT Herb Hernandez
Naval Postgraduate School, US

MAJ Sebastian Scheibe
German Army

INTRODUCTION

In 2005 the Office of Naval Research (ONR) released Broad Agency Announcement (BAA) 05-020 detailing the desired capabilities for the Transformable Craft (T-Craft). The T-Craft will provide a "game changing" capability for the US Navy's sea basing concept. T-Craft will advance the concepts of operational maneuver from the sea (OMFTS) and ship-to-objective maneuver (STOM). The T-Craft will improve the current US Navy capabilities by improving the cargo limitations of the Landing Craft Air Cushion (LCAC) and the speed limitations of the Landing Craft Unit (LCU).

The pertinent T-Craft capabilities modeled are:

1. Un-refueled range, in a no cargo condition, of 2,500 nautical miles (20 knots)
2. Full load condition speed of 40 knots
3. Amphibious capability to traverse sand bars and mud flats providing "feet dry on the beach"
4. Un-refueled range in high speed of 500-600 nautical miles (40 knots)

BAA 05-020 details the full list of desired capabilities of the T-Craft. This IDFW 20 study was conducted to examine those capabilities, while focusing on the following areas:

1. Determine critical factors and their threshold values and sensitivities.
2. Model T-Craft behavior and survivability in a hostile environment.

MODEL 1

Robust Analysis of Desired Capabilities of the Transformable Craft in Seabasing Missions

Model Description

The first model was created in Arena to simulate the transportation of troops to shore aboard T-Craft. Figure 1 provides an overview of the model. The T-Crafts loads troops at the sea base, transits to the shore, converts to Air Cushion Vehicle (ACV) mode, offloads the troops at the shore, converts back to Surface Effect Ship (SES) mode and transits back to the sea base. If necessary the T-Crafts are refueled and loaded again until all troops are projected to shore. During transit and unloading the T-Crafts may suffer from enemy hits and sink. These hits reduce the amount of troops that reach the shore and causes T-Craft loses.

Design of Experiments

Table 1 presents the input parameters used in the model. Decision factors are directly related to the desired capabilities of the T-Craft. Noise factors consider operational dependencies.

The experiment utilized a Nearly Orthogonal Latin Hypercube (NOLH) design with 29 factors and 28 rotations for all non-binary variables and a Hamardard matrix for the two binary variables (11 and 31) were crossed. The results of

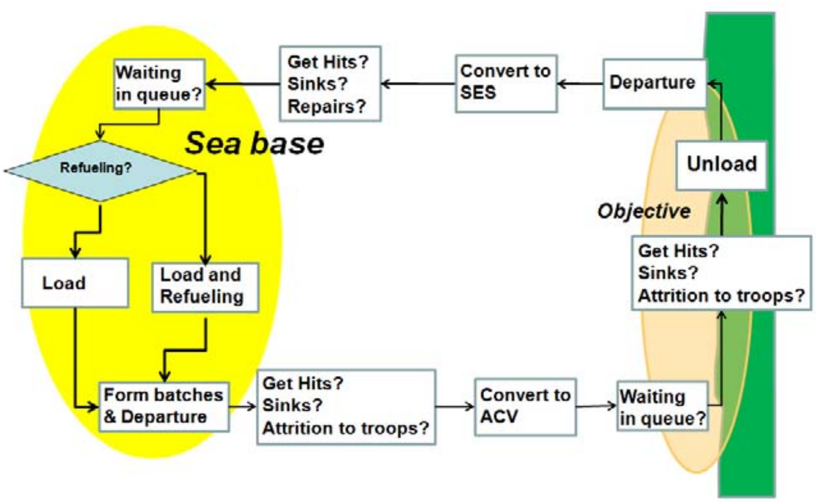


Figure 1: Generic Model Overview of the Arena Model

the crossing is 29,700 runs. For every run 4 replications were done. This effort resulted in a final total of 118,800 runs. Effectiveness was evaluated with three different Measures of Effectiveness (MOEs): the time to complete a mission (hours), the Cargo Onshore Rate, and the T-Craft Destroyed Rate.

S/N	Factor	MIN VALUE	MAX VALUE	
1	Cargo Payload Weight (LT)	300	1000	
2	Cargo Deck Size (sqft)	2200	10000	
3	Speed SES (knts)	35	55	
4	Speed ACV (knts)	4	10	
5	Load Time (hrs)	1	5	
6	Unload Time (hrs)	1	4	
7	Time to Convert to ACV (hrs)	0.25	1	
8	Time to Convert to SES (hrs)	0.25	1	
9	# T-Crafts	6	22	Decision Factors
10	# SeaSpots	5	13	
11	Refueling during loading? (1-yes, 0-no)	0	1	
12	Refueling rate (tons/hour)	80	160	
13	Tank capacity (LT)	110	150	
14	Batchsize	1	22	
15	ProbFailure	0.01	0.1	
16	Time to repair (hrs)	1	3	
17	# hits to repair	1	5	
18	Total Load (LT)	5000	40000	
19	Footprint MEB (sqft)	50000	120000	
20	Deck Use Efficiency (between 0 and 1)	0.7	0.95	
21	Distance (nm)	25	250	
22	TransitionDistance (nm)	0.5	5	
23	# ShoreSpots	3	22	
24	ProbHit Batch (%)	0.01	0.5	Noise Factors
25	ProbHit Batch during Unloading (per hour)	0.01	0.5	
26	AttritionRate Troops	0.01	0.5	
27	ProbSink (1 Hit)	0.01	0.3	
28	ProbSink (2 Hits)	0.31	0.5	
29	ProbSink (3 Hits)	0.51	0.75	
30	ProbSink (4 Hits)	0.76	0.99	
31	Distribution (0-TRIA, 1-ShiftedEXPO)	0	1	

Table 1: Decision Factors and Noise Factors

Data Analysis

Data was summarized by averaging over the noise factor space. A model was then fitted by including two-way-interactions and quadratic effects. Figure 2 shows the distribution of the MOEs.

Figures 3, 4, and 5 show the fitted model parameters for the individual MOEs and the Prediction profilers.

Figure 3 shows that the # T-Crafts, Load Time, Cargo Payload Weight, Batch Size, and Unload Time are the most significant factors affecting the mean of Time to Complete.

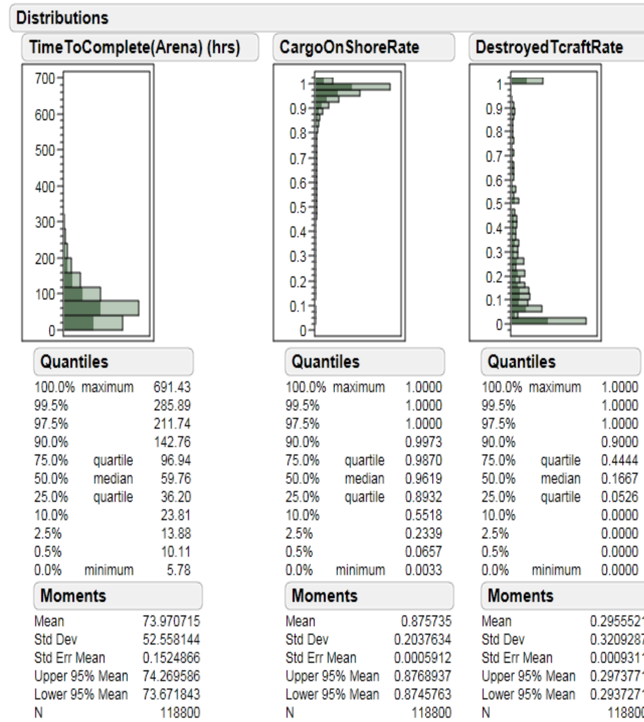


Figure 2: Distribution of the Time to Complete, Cargo Onshore Rate, and destroyed T-Craft Rate

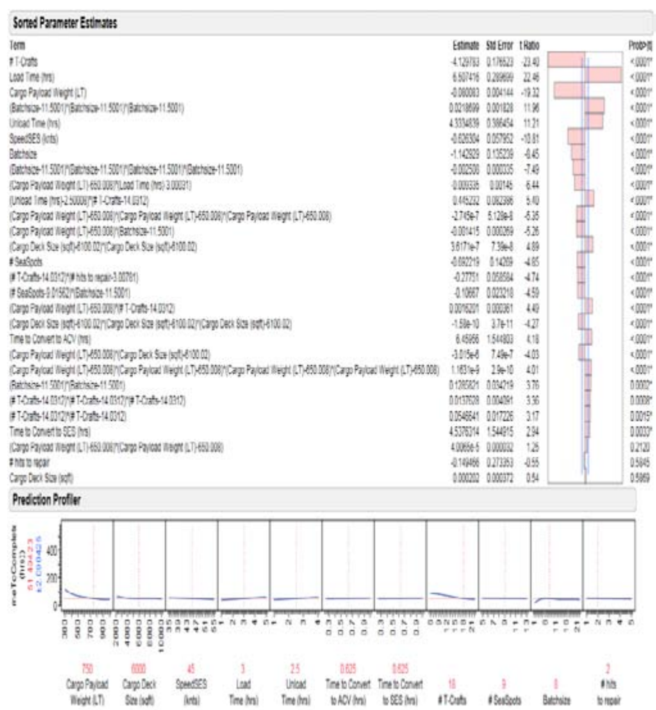


Figure 3: Sorted Parameter Estimates and Prediction Profiler of the Mean (Time To Complete) for the decision factors

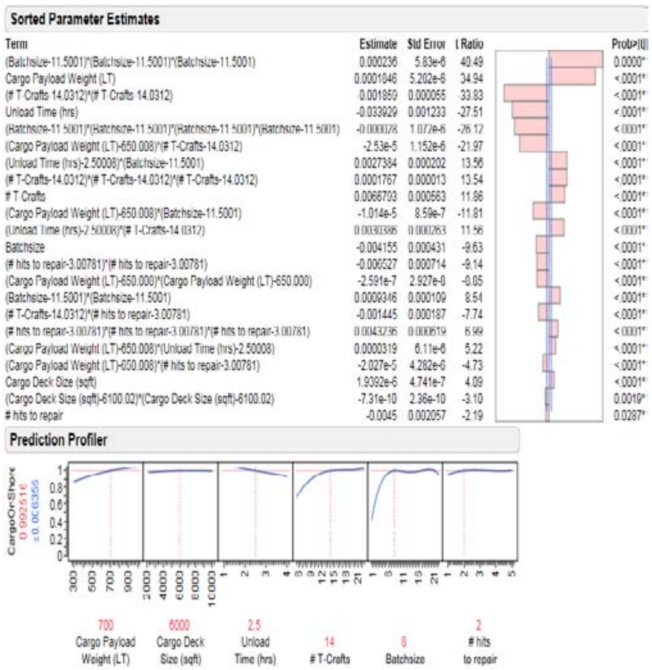


Figure 4: Sorted Parameter Estimates and Prediction Profiler of the Mean (Cargo Onshore Rate) for the decision factors

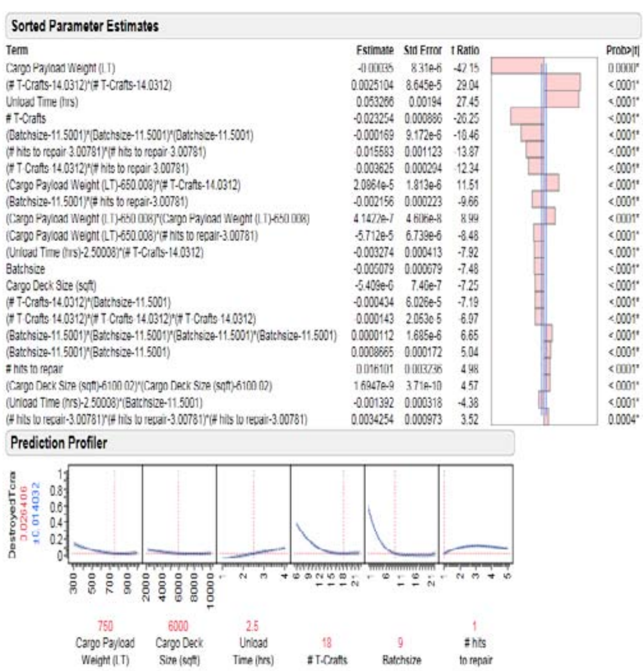


Figure 5: Sorted Parameter Estimates and Prediction Profiler of the Mean (Destroyed T-Craft Rate) for the decision factors

Figure 4 shows that Batch Size, Cargo Payload Weight, # T-Crafts, Unload Time, and Number of Hits to Repair are the most significant factors affecting the mean of Cargo Onshore Rate.

Figure 5 shows that Cargo Payload Weight, # T-Crafts, Unload Time, and Batch Size are the most significant factors affecting the mean of T-Craft Destroyed Rate.

It should also be noted that further analysis demonstrated RSquare values between 0.351 and 0.546.

Fitting partition trees on the means of the three MOEs also gave # T-Craft, Batch Size, and Cargo Payload Weight as the most significant factors.

MODEL 1 CONCLUSIONS

The most significant factors across all MOEs were: # T-Craft, Batch Size, Cargo Payload Weight, Unload Time, Load Time and Number of Hits to Repair. In order to decrease the mean Time to Completion, the number of T-Crafts should be increased. Cargo Payload Weight can also be increased. It is recommended that the number of T-Craft be maximized (within budget) and the Cargo Payload Weight should be as large as possible. The robust analysis gave us also some interesting threshold values: In order to achieve the shortest mission durations, the number of T-Craft should be at least 18 and the Cargo Payload Weight should exceed 750 LT. To get a high Cargo Onshore rate and a low destroyed T-Craft Rate, the batch size should exceed 9 and the survivability of the T-Crafts should allow two hits before major repairs are needed. Also, Load time and Unload Time should be as small as possible, and the Deck size area of the T-Craft should exceed the objective of 5500 sqft.

MODEL 2

Developing and Data Farming a Mission Model of the Transformable Craft in an Operational Environment

A second model was developed to address two major questions:

1. Does the T-Craft need an organic self-defense capability?
2. How should the T-Craft be employed when a threat exists?

The above questions were addressed with an agent based simulation tool developed by the New Zealand defense forces called Map Aware Non-uniform Automata (MANA). MANA was utilized to explore two scenarios, a peace keeping/peace enforcement operation in Columbia and a regional conflict in Malaysia. The scenarios were based on a threat assessment within the regions. The models assume that the sea base consists of two amphibious ready groups (ARG), section of T-Craft, LCS surface and anti-surface packages, and MPF(F).

Design of Experiments

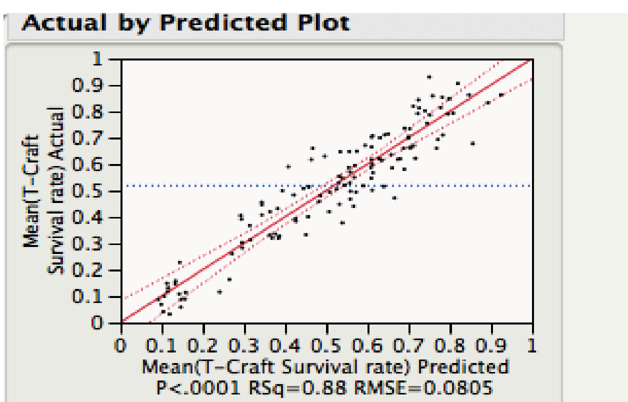
The initial DOE was focused on gleaning insight on how T-Craft survivability is affected in four different cases: armed, not armed, armed and escorted, and not armed and escorted. The design consisted of a NOLH with 8 factors and 125 design points. Each design was run 50 times. The measure of effectiveness in the experiment is the mean T-Craft survival rate. The factors are their levels are summarized in Table 2.

Factor Name	Range
Number T-Craft	5 – 11
T-Craft Armed	1 – 0 (Yes – No)
T-Craft Escort	1 – 0 (Yes – No)
Alt Waypoint (back to Sea Base)	1 – 0 (Yes – No)
Speed	35 – 55
Number Red Patrol Boats	5 – 10
Number Red Semi-submersibles (scenario I)	5 – 10
Number Red SWARM boats (scenario 2)	5 – 10

Table 2: Factors and Ranges

Data Analysis

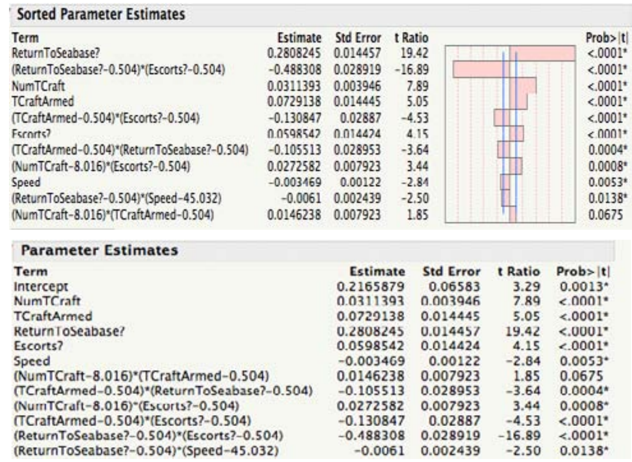
After the initial runs with both models, escorting and arming the T-Craft appeared to have increased survivability. The results were obtained through regression analysis, partition plots, and analysis of the distribution. The regression analysis and partition plot results for the first scenario (Columbia) are presented in Figures 6 and 7. The same results for the second scenario (Malaysia) are presented in Figures 8 and 9.



Summary of Fit

RSquare	0.883089
RSquare Adj	0.871708
Root Mean Square Error	0.080531
Mean of Response	0.518064
Observations (or Sum Wgts)	125

Figure 6: Scenario 1 Regression Analysis



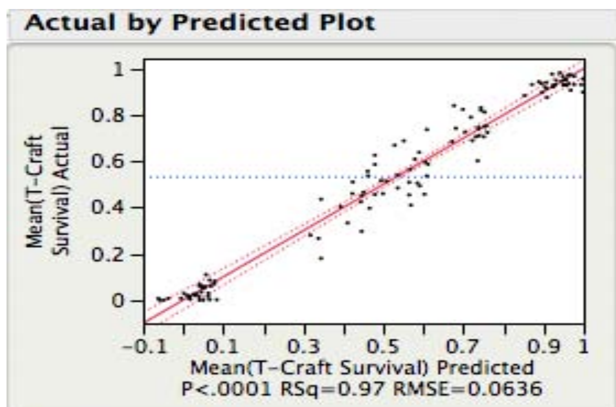
Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
ReturnToSeabase?	0.2808245	0.014457	19.42	<.0001*
(ReturnToSeabase?-0.504)*(Escorts?-0.504)	-0.488308	0.028919	-16.89	<.0001*
NumTCraft	0.0311393	0.003946	7.89	<.0001*
TCraftArmed	0.0729138	0.014445	5.05	<.0001*
(TCraftArmed-0.504)*(Escorts?-0.504)	-0.130847	0.02887	-4.53	<.0001*
Escorts?	0.0598542	0.014424	4.15	<.0001*
(TCraftArmed-0.504)*(ReturnToSeabase?-0.504)	-0.105513	0.028953	-3.64	0.0004*
(NumTCraft-8.016)*(Escorts?-0.504)	0.0272582	0.007923	3.44	0.0008*
Speed	-0.003469	0.00122	-2.84	0.0053*
(ReturnToSeabase?-0.504)*(Speed-45.032)	-0.0061	0.002439	-2.50	0.0138*
(NumTCraft-8.016)*(TCraftArmed-0.504)	0.0146238	0.007923	1.85	0.0675

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	0.2165879	0.06583	3.29	0.0013*
NumTCraft	0.0311393	0.003946	7.89	<.0001*
TCraftArmed	0.0729138	0.014445	5.05	<.0001*
ReturnToSeabase?	0.2808245	0.014457	19.42	<.0001*
Escorts?	0.0598542	0.014424	4.15	<.0001*
Speed	-0.003469	0.00122	-2.84	0.0053*
(NumTCraft-8.016)*(TCraftArmed-0.504)	0.0146238	0.007923	1.85	0.0675
(TCraftArmed-0.504)*(ReturnToSeabase?-0.504)	-0.105513	0.028953	-3.64	0.0004*
(NumTCraft-8.016)*(Escorts?-0.504)	0.0272582	0.007923	3.44	0.0008*
(TCraftArmed-0.504)*(Escorts?-0.504)	-0.130847	0.02887	-4.53	<.0001*
(ReturnToSeabase?-0.504)*(Escorts?-0.504)	-0.488308	0.028919	-16.89	<.0001*
(ReturnToSeabase?-0.504)*(Speed-45.032)	-0.0061	0.002439	-2.50	0.0138*

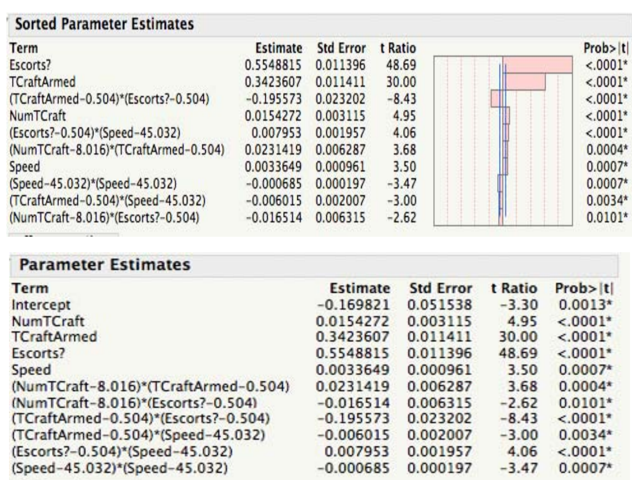
Figure 7: Scenario 1 Partition Plots



Summary of Fit

RSquare	0.969377
RSquare Adj	0.966691
Root Mean Square Error	0.06362
Mean of Response	0.531346
Observations (or Sum Wgts)	125

Figure 8: Scenario 2 Regression Analysis



Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Escorts?	0.5548815	0.011396	48.69	<.0001*
TCraftArmed	0.3423607	0.011411	30.00	<.0001*
(TCraftArmed-0.504)*(Escorts?-0.504)	-0.195573	0.023202	-8.43	<.0001*
NumTCraft	0.0154272	0.003115	4.95	<.0001*
(Escorts?-0.504)*(Speed-45.032)	0.007953	0.001957	4.06	<.0001*
(NumTCraft-8.016)*(TCraftArmed-0.504)	0.0231419	0.006287	3.68	0.0004*
Speed	0.0033649	0.000961	3.50	0.0007*
(Speed-45.032)*(Speed-45.032)	-0.000685	0.000197	-3.47	0.0007*
(TCraftArmed-0.504)*(Speed-45.032)	-0.006015	0.002007	-3.00	0.0034*
(NumTCraft-8.016)*(Escorts?-0.504)	-0.016514	0.006315	-2.62	0.0101*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-0.169821	0.051538	-3.30	0.0013*
NumTCraft	0.0154272	0.003115	4.95	<.0001*
TCraftArmed	0.3423607	0.011411	30.00	<.0001*
Escorts?	0.5548815	0.011396	48.69	<.0001*
Speed	0.0033649	0.000961	3.50	0.0007*
(NumTCraft-8.016)*(TCraftArmed-0.504)	0.0231419	0.006287	3.68	0.0004*
(NumTCraft-8.016)*(Escorts?-0.504)	-0.016514	0.006315	-2.62	0.0101*
(TCraftArmed-0.504)*(Escorts?-0.504)	-0.195573	0.023202	-8.43	<.0001*
(TCraftArmed-0.504)*(Speed-45.032)	-0.006015	0.002007	-3.00	0.0034*
(Escorts?-0.504)*(Speed-45.032)	0.007953	0.001957	4.06	<.0001*
(Speed-45.032)*(Speed-45.032)	-0.000685	0.000197	-3.47	0.0007*

Figure 9: Scenario 2 Partition Plots

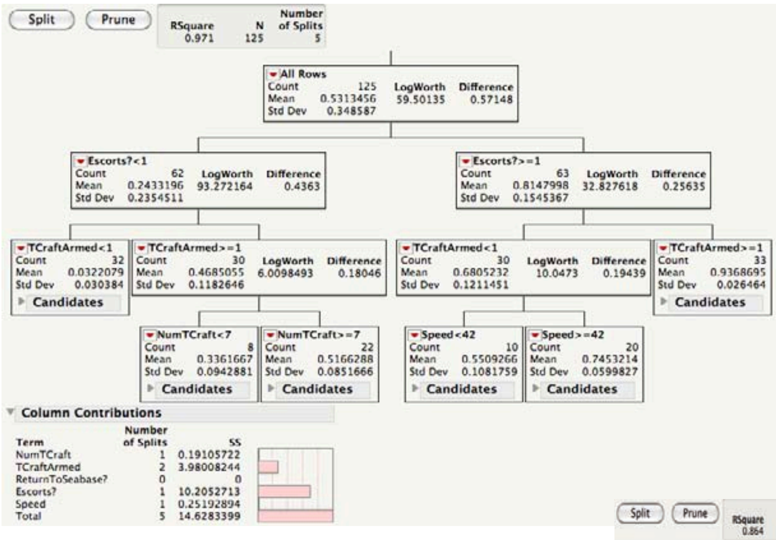


Figure 11: Scenario 2 Analysis of Distribution

The data demonstrates that, for the short range scenario, the TTP return to seabase when enemy are present appeared to be significant with respect to the T-Craft survival rate. This factor is followed by escorts and speed. In the Malaysia scenario, escort and armed appeared to be significant (in that order). Further, speed appeared to have an effect on T-Craft survivability. The distributions were analyzed to confirm the conclusions. Figure 10 presents the Columbia scenario and Figure 11 presents the Malaysia scenario.

The data presented in the analysis of distribution diagrams confirms the previous conclusions.

MODEL 2 CONCLUSIONS

The IDFW allowed for the refinement of the scenarios in MANA. The analysis of the model demonstrated that escorts and arming the T-Craft appear to be significant factors. Potential future work should include expanding the model to accommodate future design runs that vary the weapon systems on the T-Craft. This should provide insight into what escort and weapon mix is optimum for T-Craft survivability. Further expansion should also include more layers of the threat and friendly posture. Counter measures and radar jamming equipment could also be added to the T-Craft to see how that affects survivability.

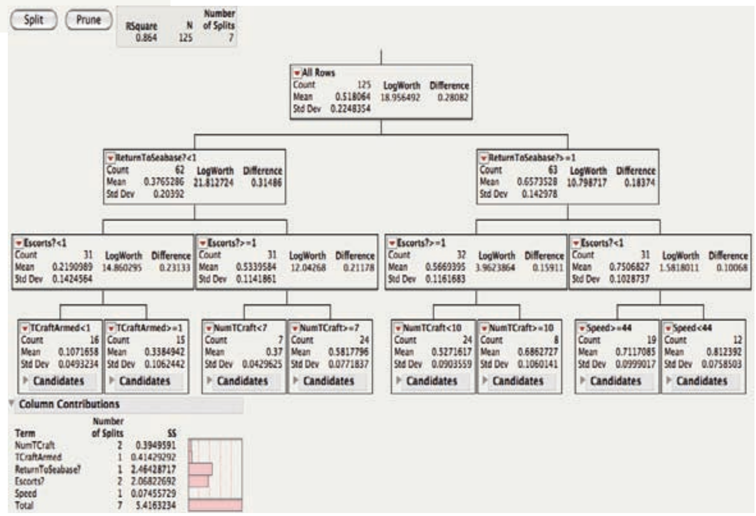


Figure 10: Scenario 1 Analysis of Distribution

Team 3: Using Data Farming to Develop and Evaluate IED Scenarios for PAX

TEAM 3 MEMBERS

SEICHTER, Stephan LtCol
Bundeswehr Transformation Center, Germany

LAMPE, Thorsten
MALY, Andreas
EADS System Design Centre, Germany

ANG, Chee Chien
DSO National Laboratories, Singapore

THOMPSON, Meredith
NPS, SEED Center for Data Farming, USA

WHITNEY, Laura
NPS, SEED Center for Data Farming, USA

INTRODUCTION

The Bundeswehr Transformation Center is examining how M&S can effectively support Concept Development and Experimentation (CD&E) projects analyzing different aspects of stabilization operations. Human Factors and Human Behavior analyses have been highly relevant in this context. Current studies examine the use of the agent-based model PAX for this purpose. These studies include Humanitarian Assistance scenarios in response to disaster-caused refugee movements and Irregular Warfare scenarios modeling asymmetric tactics instead of force-on-force engagements. The analysis focuses on situations in which the forces are facing adversaries using improvised weapons, namely Improvised Explosive Devices (IEDs).

Scenario

The current scenario being analyzed with the model PAX takes place inside a refugee camp operated by the German military. The scenario has 60 civilian agents of the same ethnic group; therefore, we assume that there are no acts of aggression among the civilians themselves.

Because the current focus of the analysis is to develop a realistic behavior for the civilians in the refugee camp looking to the possible outcomes such weapons would have on civilian populations, the scenario does not contain any military personnel yet, even though the camp is assumed to be run by the military.

At the beginning of the scenario an IED detonates among a group of civilians. The explosion harms several agents standing close to the IED, while other agents residing outside the area of the detonation do not suffer any injury.

The further course of the simulation is then controlled by the direct and indirect effects of the explosion on the civilians' physical as well as emotional conditions.

Figure 1 depicts the initial setup of the refugee camp scenario that was used as the baseline for the studies performed during the week of the workshop.

Questions

In preparation for the workshop, major adjustments had been made to integrate civilian behavior in response to an IED explosion into the model PAX. As a result, the team was able to work with the model itself during the workshop, and view and review the current status of the implementation and scenario.

Among the effects that had already been considered and integrated into the model were

- possible physical damage caused by the IED, depending on the actual size of the IED
- emergence of fear among the agents who witness the detonation
- emergence of fear among the agents who recognize the detonation's effects in wounding or killing fellow civilians
- emergence of curiosity among agents once their fear has subsided
- emergence of a motivation of helpfulness towards injured agents after the detonation

The primary goal during IDFW20 was to use Data Farming and single run analysis to evaluate and further develop the current IED scenario in order to more accurately

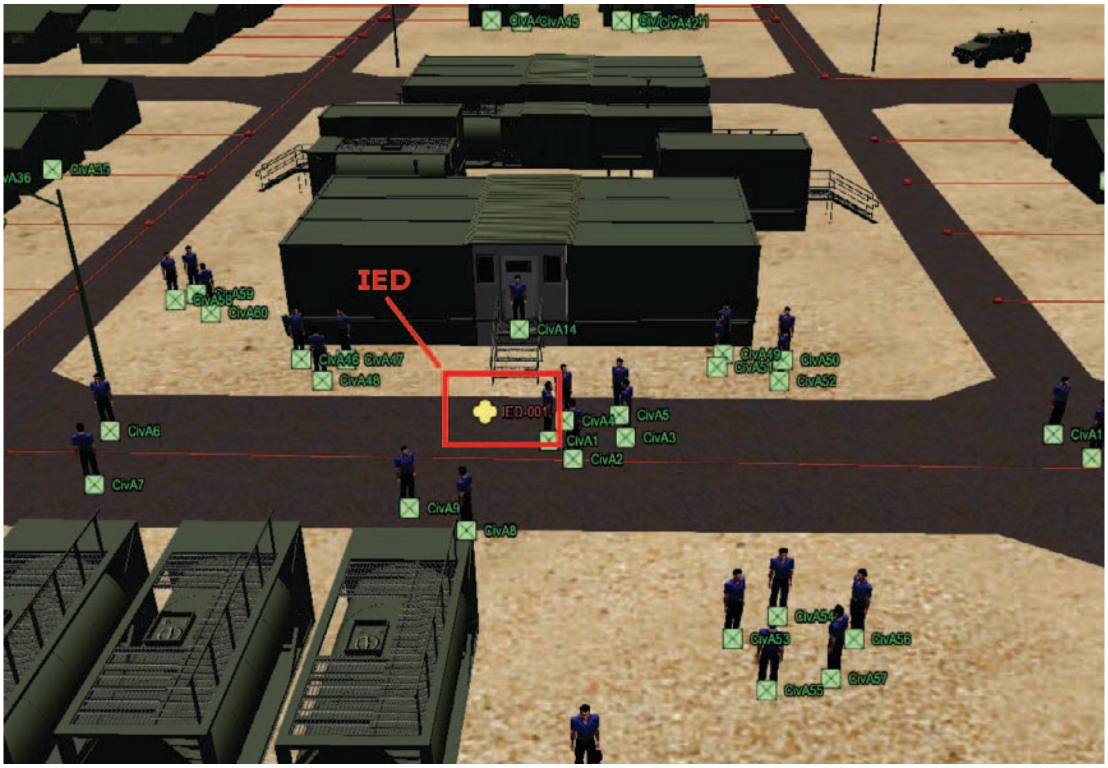


Figure 1: Initial setup of the scenario; the position of the IED among the civilian group is marked

model realistic human behavior and comprehensible actions and reactions to IED explosions.

The workgroup was driven by guiding questions such as:

- What effects do IED explosions have on human behavior?
- Have the relevant factors and elements characterizing such a situation been identified and captured in the current modeling of the agents' behavior?
- Can an IED detonation lead to aggressive behavior among the civilians? If yes, what are the reasons and when does aggression start to dominate the situation?

STUDIES & ANALYSES

As a first step, we investigated the effects that differently-sized IEDs had on the surrounding agents. This effort also served as verification for the new implementation of IEDs in PAX and the effects of detonations on human agents, as we were able to verify that different sizes of IEDs do indeed lead to a respective increase or decrease in injury and death within the model.

Experiment 1 Setup

In the model, the damage caused by an IED explosion is expressed in so-called "kill levels". Every kill level defines a type of damage it can cause, a probability with which this damage is actually caused, and a diameter around the center of the detonation inside which the kill level can cause that damage. The diameters (defined in meters) of the kill levels determine their mean areas of effectiveness (MAEs).

In this experiment we distinguished three types of kill level MAEs:

- MAE_CompleteKill: Agents inside this area can be killed by the detonation.
- MAE_MobilityKill: Agents inside this area can be so heavily injured by the explosion that they are unable to move afterwards.
- MAE_GeneralKills: Agents inside this area can suffer injuries that reduce their overall health, but do not kill them or leave them immobilized.

For our experiments, we generated IEDs of different sizes by modifying their respective MAEs. Table 1 shows the ranges we varied for each IED size. The resulting ten different sizes of IEDs were consecutively numbered "IED size 1" to "IED size 10". For simplification, variation of an IED's damaging power was done solely by this modification of the MAEs, while the kill probability inside every MAE was fixed to a value of 80% throughout all our experiments.

MAE_ CompleteKill (diameter)	MAE_ MobilityKill (diameter)	MAE_ GeneralKills (diameter)	IED size
2	3	4	1
4	6	8	2
…	…	…	…
20	30	40	10

Table 1: Definition of IED sizes by modification of the MAE diameters of the different kill levels

Figure 2 shows a visual representation of the various kill levels and their MAEs for the largest IED that was used in our experiments.

Figure 2: Visual representation of kill levels and MAEs of IED size 10 ("Large IED")

Experiment 1 Analysis

In examining the results of the first Data Farming experiment, we found that an increase in IED size led to an almost linear increase in the number of dead and injured people, as was expected. The mean number of dead and injured (heavily injured or immobilized) people (MOE_NumVictims) varied from 3.5 to 18.3 depending on the IED size.

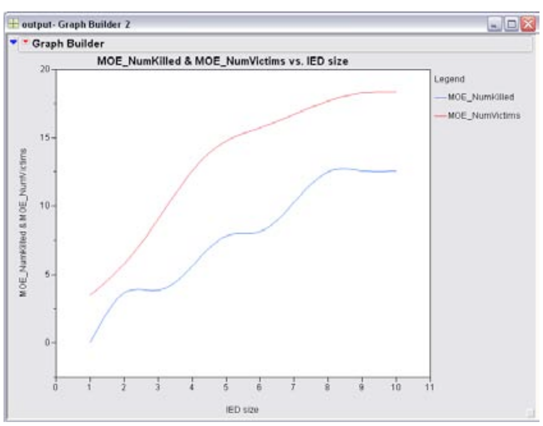


Figure 3: Correlation of IED size and MOE_NumKilled and MOE_NumVictims

Furthermore, we verified that inside the various MAEs indeed about 80% of the agents suffered the respective kill level, as was intended.

Based on these results the team regarded the implementation of IEDs in the model as correct and the effects caused by the detonation as plausible. Out of the 10 defined IED sizes, three were selected, labeled "Small IED" (IED size 1), "Medium IED" (IED size 4) and "Large IED" (IED size 10) in the following experiments. For ease of analysis these three IED sizes were once again renamed to "IED size 1" (Small IED) through "IED size 3" (Large IED).

Experiment 2 Setup

Having determined the IED sizes to be used from experiment 1, this experiment directly targeted the questions formulated above. We wanted to see what effects the IED detonation would have on agent behavior and if the effects seemed plausible to the participants of the working group.

To examine these questions, we set up an NOLH design in which we varied the following parameters:

- *Sensor Range:* The distance in meters up to which an agent can perceive dead and injured agents in his proximity.
- *Personality Constant "Rise Of Aggression":* Determines an agent's tendency to act aggressively in reaction to perceiving dead and injured people.
- *Personality Constant "Rise Of Anger":* Determines an agent's tendency to become angry in reaction to perceiving dead and injured people.
- *Personality Constant "Rise Of Helpfulness":* Determines an agent's tendency to become helpful if he encounters people in need of help in his proximity.
- *Duration of a Rescue Action:* Helpful agents help those in need by rescuing them and bringing them to the medical facilities located in the refugee camp. This parameter defines the mean duration of such a rescue action.

Table 2 shows the ranges in which the parameters were varied.

Parameter	Min	Max
Sensor Range	10.0	70.0
Constant Rise of Aggression	0.0	1.0
Constant Rise of Anger	0.0	1.0
Constant Rise of Helpfulness	0.0	1.0
Duration of a Rescue Action	120	600

Table 2: Variation of parameters in experiment 2

The Measures of Effectiveness (MOEs) we wanted to investigate were the following:

- *MOE_Time_60_pc_Angry:* The duration of time it takes for 60% of the civilian population to get angry (i.e. the percentage of agents whose action-leading motive is anger). If that anger level is reached in the population, we assume that anger is "dominating the situation" as referred to in the questions posted above.
- *MOE_NumRescuedCivilians:* The number of civilians that have been rescued by the end of the scenario.
- *MOE_PercentageRescuedOutOfVictims:* The final percentage of rescued civilians in relation to the number of people who actually need to be rescued. People in need of rescue are called "victims" here, encompassing all heavily injured agents, i.e. immobilized.
- *MOE_FinalAverageFear:* The average fear level in the civilian population at the end of the simulation.
- *MOE_FinalAverageAnger:* The average anger level in the civilian population at the end of the simulation.
- *MOE_NumCiviliansHelpfulness:* The number of agents who, at the end of the simulation, are willing to help other agents.

Experiment 2 Analysis

Two quite unexpected observations were that the point in time when 60% of the population was led by their anger was reached very quickly, and that very few of the victims actually received the help they needed. The reasons were found to be threefold:

- The range of the parameter variation in this experiment was, intentionally, very wide in order to get a coverage of the whole parameter spectrum. This situation also meant that lots of rather improbable, but still feasible combinations of parameters were included in the results, such as agents getting very angry due to the perception of victims in their proximity combined with not feeling any urge to help these perceived victims.
- The perception of the detonation as such leads to an instant increase of anger in the agents. This effect in the model seems to be quite dominant within the range in which the anger personality constant parameter had been varied, and therefore anger takes over as the leading motive in many of the simulated scenarios instantly after the detonation.
- We assumed that when agents perceive victims, they also react with an increase of anger as a reaction to the fact that no help is provided for those in need. The model formula responsible for this reaction generated an exponential increase in anger (the longer untreated victims are perceived, the quicker the anger rises), which resulted in many of those agents standing close to victims becoming angry very quickly.

The combination of these three factors led to anger becoming the dominating factor too quickly in the scenario, thereby not allowing the helpfulness characteristic to ever dominate. As a result, few victims were rescued.

The team participants deemed this course of the scenario rather unrealistic. The extremely rapid increase of anger due to the perception of untended victims seemed especially improbable.

Experiment 3 Setup

As a result of the findings of experiment 2, two main changes were made.

First, as proposed by the workgroup, the agents' progression of anger due to the perception of untended victims is no longer an exponential increase, but rather a linear increase instead. This change still leads to people getting angry from the sight of wounded and dead people, but, as judged by the workgroup, in a more plausible way.

Second, the parameter ranges were adapted to preclude parameter combinations in which people become very angry but do not become helpful. Table 3 shows the adapted parameter ranges that were used in experiment 3.

Parameter	Min	Max
Sensor Range	40.0	80.0
Constant Rise of Aggression	0.0	0.6
Constant Rise of Anger	0.0	0.6
Constant Rise of Helpfulness	0.5	2.0
Duration of a Rescue Action	120	600

Table 3: Parameter ranges in experiment 3

Experiment 3 Analysis

As expected, the increase of the agents' anger was reduced. However, the behavior of the agents still did not meet the workgroup's expectations, as still only very few of the victims received the help they needed.

As opposed to being occupied by their anger, the agents in this experiment were mainly restrained from helping due to the fact that fear was overwhelming their other personality constants. One of the main problems was that agents who moved to the site of the detonation due to curiosity stayed away too far from the point of the actual incident (curious agents keep their distance from the place of the detonation for the sake of their own security). On the other hand, the model formula responsible for calculating the helpfulness of an agent A towards a victim B heavily depends on the proximity of A to B. We therefore assumed that these two distance factors collided with each other.

Experiment 4 Setup

To investigate if our assumption was correct and if we could indeed get better results by merely having the curious agents move in closer to the victims to allow them to develop a helpfulness motivation, we extended experiment 3 with a follow-up experiment.

The only factor actually changed in the setup was the distance curious agents keep to the site of the detonation. All other parameters remained unchanged.

Experiment 4 Analysis

The analysis of this fourth experiment showed that the model produced reasonable results regarding the expected course of the scenario, at least in some of the cases, if we brought the potential helpers closer to the victims, thereby helping them to generate a higher motivation of helpfulness towards the wounded.

By changing the single parameter from experiment 3 to experiment 4 we eliminated the dominance of the proximity factor built into the model's helpfulness formula. However, it came to a surprise to us that this change only led to the expected results in some of the cases, not in all of them.

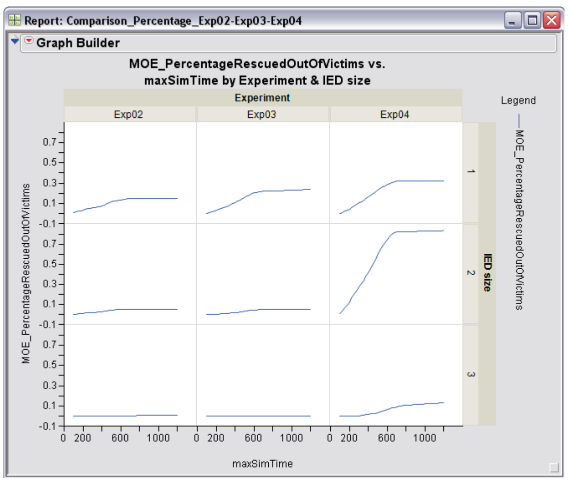


Figure 4: Percentage of rescued victims of the course of time; comparison of Exp 2 - 4 and IED sizes 1 - 3

Figure 4 shows the percentage of rescued victims over the course of the simulation time. Obviously all graphs are monotonously increasing over simulation time. We indeed did receive a significant change between experiment 3 and experiment 4, but only in the case of the medium-sized IED. In the cases of small IEDs and large IEDs, the latest changes did not yield any significant improvements in terms of people being rescued.

It has to be examined in follow-on experiments why the changes in the scenario yielded such good results just for one particular IED size.

CONCLUSION

The primary goal of this workgroup was to evaluate and further develop the current IED scenario of PAX and to face validate the new functionalities that had been introduced into the model for simulating the effects of an IED detonation on the civilian population of a refugee camp. We clearly achieved this goal and were able to gainfully use the Data Farming methodology and its power to explore the parameter space for scenario evaluation and development by subsequently refining our scenario and adjusting the parameter ranges.

Before we started our work, the workgroup's expectations regarding the effects of an IED detonation on the civilian population inside such a camp were captured. It turned out that the effects and factors that had been identified as important during the development of the model matched the expectations of the workgroup to a high degree.

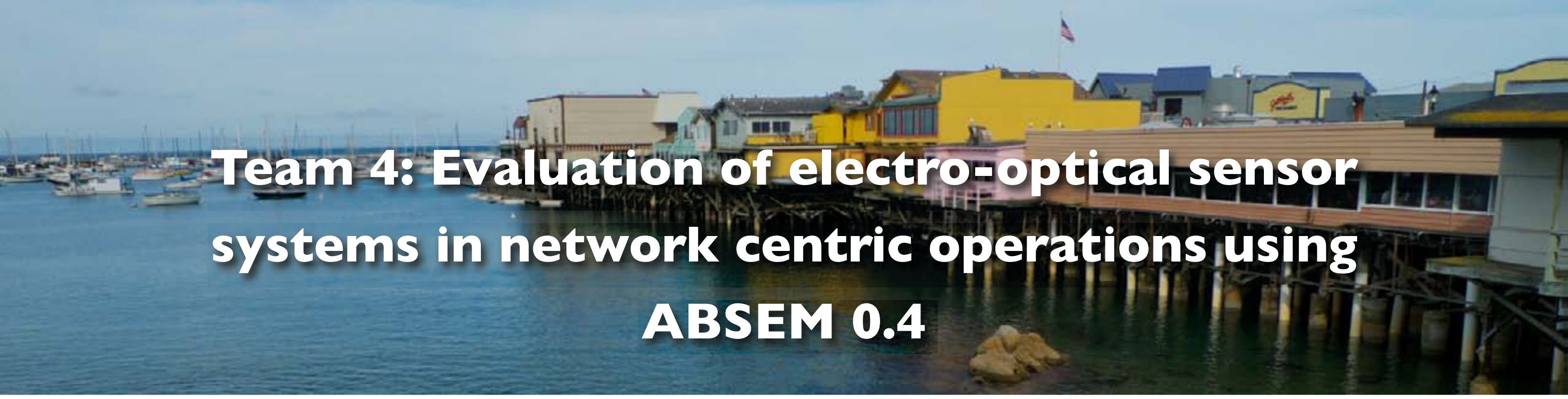
Furthermore, by commonly analyzing our experiments and discussing the findings, we were able to, step by step, modify the scenario as well as certain aspects of the model itself during the week. Therefore we were able to generate more realistic courses of action in our scenario.

Although we made good progress during IDFW20, we were not able to attend to all the aspects that were mentioned in our discussions. But we gained valuable insights that enabled us to derive further studies of modeling the effects of IED attacks on civilian populations. We recommend futher investigations looking to the following:

- The distance factors contained in the model formulas for curiosity and especially helpfulness seem to be too influential. It could be considered whether the proximity to the victims should play any role at all in terms of people's motivation to help.
- At the moment there is no connection in the model between the elements "anger arising due to the perception of victims" and "helpfulness arising due to the perception of victims". We recognized that an agent who perceives a victim becomes angry over the fact that the victim is unattended, but does not try to help himself. This combination might be a possible behavior in such a situation. But it should also be possible that although being angry, an agent helps a victim. We will attempt to adapt the model in order to represent either behavior.
- The effect that an explosion instantly increases the anger of people witnessing the event needs to be reviewed and validated by empirical findings from the scientific field of psychology.
- The discussions during the week furthermore generated some new ideas for additional civilian behavior as a reaction to the detonation. Among those were suggestions such as panic or shock reactions, which are not yet part of the implemented civilian behavior.

These aspects will be addressed in the further development process of the model and the refinement of the IED scenario.

Finally it has to be said that during the four days of IDFW20 we had a great team and an extremely ambitious and collaborative atmosphere to work in. Every member of the workgroup contributed to the development of the model and the scenario with valuable comments, suggestions and contributions during our fruitful discussion. Special thanks to all members of Team 03 – THE BEST!.



Team 4: Evaluation of electro-optical sensor systems in network centric operations using ABSEM 0.4

TEAM 4 MEMBERS

André Geiger - Lead
Dr. Daniel Nitsch
Federal Office of Defense Technology and Procurement, Germany

OTL Thomas Erlenbruch
Army Office, Germany

Choo Chwee Seng
DSO, Singapore

Ka-Yoon Wong
Naval Postgraduate School, USA

Dr. Klaus-Peter Schwierz
Daniel Kallfass
Gudrun Wagner
EADS, Germany

INTRODUCTION

After 2 years of intensive development, the new ABSEM version 0.4 was released and presented at IDFW20 to the International Data Farming Community. The model concentrates on modeling complex technical aspects in NCO and to do so, it integrates detailed physical theories when it comes to simulating the output of various sensors and when determining the effect of different weapon systems. The very realistic and physically based modeling allows for getting reliable answers to posed questions regarding specific scenarios. This capability could again be confirmed when working with the model during IDFW20 looking at an ambush scenario located in Mazar-e-Sharif.

Objectives

In Data Farming experiments the team's main intention was to examine the effect of different patrol compositions (i.e. different type and number of vehicles along with their corresponding sensor and effector systems) under varying ambush conditions, such as different weather-dependent atmospheric conditions, varying number, positioning and equipment of insurgents.

Overall, the team had the following goals:

- Review and face validate ABSEM version 0.4
- Conduct Data Farming experiments
- Identify necessary ABSEM improvements
- Evaluate technical and tactical patrol compositions to minimize own casualties and damages; i.e.:
  - Distance of patrol vehicles
  - The benefit of deploying reconnaissance UAVs in advance
  - The effect of the patrol's reaction times (regarding processing sensor information or being ready to defend)
  - The deployment of sensor and effector systems under different circumstances

Ambush scenario

On its way from the military camp towards Mazar-e-Sharif a patrol gets ambushed by adversary insurgents equipped with rifles, anti-tank missiles and heavy machine guns. The insurgents are well camouflaged and can hardly be identified in advance. An IED triggered via remote control forces the patrol to come to a halt. The implemented patrol composition is shown in figure 1, whereas figure 2 shows a screenshot of an ABSEM simulation run.

In different scenario vignettes the goal was to analyze several ambush situations with varying force ratios and different sensor and effector systems applied by the blue and red forces.

Vehicles of type **A** and **B**	**TCT** (Tactical CIMIC Team):	**JAM** (Jammer)	**BAT** (Mobile Medic Team)	**ANP** (Afghan National Police)
Dingo (german strongly armored fighting vehicle) or Fuchs (german transportation vehicle)	Wolf (german lightly armored vehicle)	Fuchs	Fuchs	Pickup

Figure 1: Patrol composition

Data Farming Experiments

We were executing a series of data farming experiments, being interested in the influence of the following parameters:

- Blue forces variation:
  - types of vehicles: Dingo or Fox
  - distance between vehicles: from 25 meters to 100 meters in 25m steps

Figure 2: Patrol being ambushed by insurgents

- available sensor system: normal human viewing and / or infrared
- available effector system: 7,62mm machine gun, heavy machine gun
- rules of engagement: from "defense only" vs. "fire as soon as insurgents were identified"
- Dingo kill probability when being hit by an RPG: 50% vs. 75%
- deployment of the UAV LUNA for advance airborne reconnaissance of the patrol route

- Red forces variation:
    - number and position of insurgents: three different ambush situations
    - available effector systems: AK47, RPG7, heavy machine gun
    - patrol vehicle that gets struck by IED (position within the patrol)
- Weather: clear sky / cloudy weather / strong fog

As MoEs we were mainly looking at damages and losses of the blue and red forces.

All of our experiments were successfully executed on the new 520-node German cluster owned by the German Procurement Office (BWB).

In an iterative approach we were executing several data farming experiments to analyze the mentioned parameters' influence on the overall mission success, i.e. avoiding blue losses. The final experiment encompassed 64800 simulation runs, successfully executed on the 520 cluster nodes in only 6 hours.

Data Farming Results

For the analyzed ambush situations, the red forces benefit a lot from bad weather conditions. In case there is strong fog, the insurgents lying in ambush and being concealed by small bushes and trees can hardly be detected, neither by the patrol itself nor by the LUNA reconnoitering the area in advance. As a result, many blue vehicles were severely damaged but only a few insurgents got killed.

For that reason we decreased the amount of data to be analyzed by only looking at the scenarios with good weather conditions.

In those cases, the type of ambush setup by the insurgents was the most influencing factor regarding blue kills, followed by the patrol vehicle distance (see Figure 3).

In ambush scenario 1, there is only one group of six insurgents south of the patrol route. They are very well hidden and camouflaged within bushes.

In ambush scenario 2, two additional groups of six insurgents are hiding behind walls and buildings north of the patrol route. They move out of their hiding places towards the patrol as soon as the IED detonates.

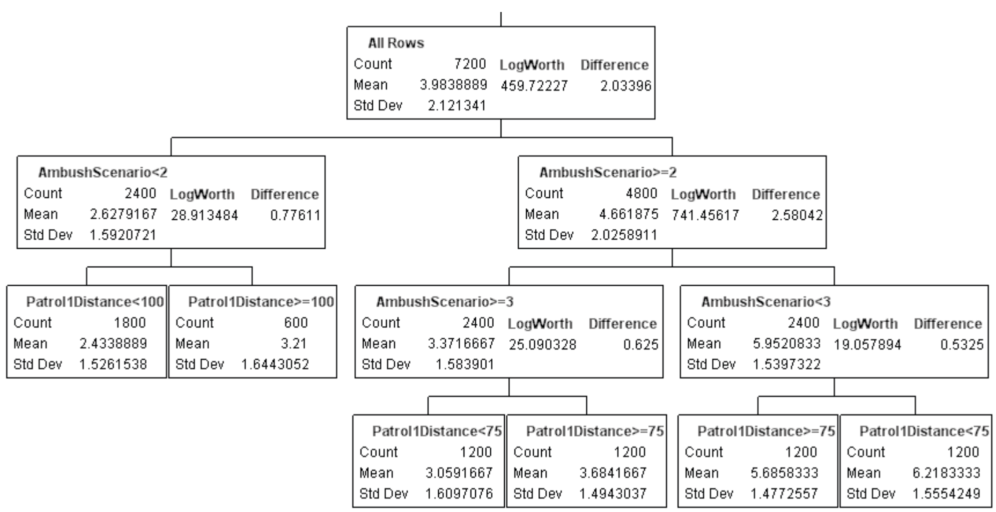


Figure 3: Regression tree for the number of blue forces destroyed

Ambush scenario 3 equals scenario 1 with an additional suicide attacker: a pickup car bomb loaded with 100kg of explosive TNT drives into one of the patrol's vehicles.

However, this suicide attacker hardly made any difference in the scenario outcome. This situation can also be seen in figure 4 showing the number of blue and red kills for the different ambush scenarios and varying vehicle distances. As expected, due to the large number of insurgents attacking the patrol, in scenario 2 the number of blue kills is rather high.

Another result is that the distance of the vehicles seems to be quite important. Short distances result in rather high blue damages , because all of the vehicles are attacked by the insurgents. But unlike what the team thought in advance, large distances may even result in worse results and more blue forces damaged. By looking at single simulation runs, we could find the following explanation: In the case of a large distance only some of the patrol vehicles are within the insurgents' effector range. So the insurgents concentrate on those few vehicles and usually then succeed in destroying those. Additionally, the patrol vehicles far away from the camouflaged insurgent group do not have a chance of attacking and destroying these insurgents.

In the experiments we also varied the number of insurgents being equipped with an RPG7 and 4 grenades each. Surprisingly, this situation was not relevant for the scenario outcome. Using the AK47 is just as good, because there is usually less dispersion and the amount of ammunition is much higher.

The same applies to the soldiers' rules of engagement. Even if the patrol was allowed to attack any identified insurgent, this situation had hardly any influence. This result is due to the following reasons. In the first place it's very hard to even detect the insurgents in time (i.e. before the IED detonates). Secondly, even if the patrol succeeds in identifying the ambush in advance, this situation more or less only changes the duration of the fight and the only targets that can be attacked are those who lie in ambush in the south (to all the others there is not even a line-of-sight).

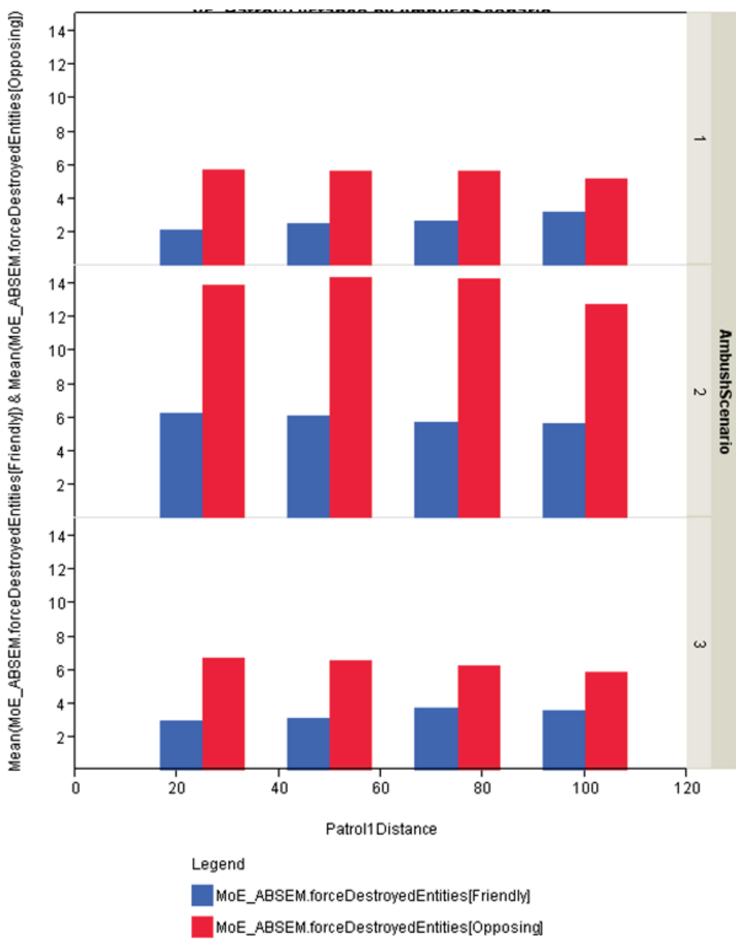


Figure 4: Number of blue and red kills for different ambush scenarios and varying patrol vehicle distances

SUMMARY AND WAY AHEAD

Once again, the intensive work with the model ABSEM confirmed the chosen approach of modeling technical systems based on quite detailed physics. This process delivers very plausible results that meet the military expectations. In combination with the very powerful 3d-visualization, the model ABSEM experiences broad acceptance by various military users.

Thanks to our team members we had many interesting discussions and got very valuable remarks for useful model enhancements.

In following activities, the ABSEM model features will be enhanced (e.g. adding the possibility of modeling tracking radar systems) and new user interfaces for setting up and analyzing complex scenarios will be implemented.

Team 5: Using Experimental Design and Data Analysis To Study The Enlisted Specialty Model For The U.S. Army G1

TEAM 5 MEMBERS

Major Robert Erdman
Captain Jeff Groves
Professor Rachel Johnson
Naval Postgraduate School, US

Colonel Kent Miller
Lieutenant Colonel Mary Lou Hall
U.S. Army

INTRODUCTION

During the International Data Farming Workshop (IDFW) 20, Team 5 worked in direct support of MAJ Erdman's thesis. MAJ Erdman's thesis work is being conducted for the Army G1, which is the branch of the Army that is in charge of all Army personnel. The G1 is responsible to develop, manage and execute all manpower and personnel plans, programs and policies – across all Army Components – for the entire Army team [1].

The Army manpower program is a 30.6 Billion dollar annual investment. Its size, diversity in the skills it needs, the cost in terms of dollars, and years to produce skilled Soldiers requires that the manpower program be closely managed. The G1 uses the Active Army Strength Forecaster (A2SF) which consists of three mathematical models to manage this manpower program. These three models are used in conjunction with one another to ensure the Army has an adequate number of people by grade and skill in order to fight the Nation's wars. One of these three models is the Enlisted Specialty (ES) model, which specifically forecasts the enlisted soldiers in the Army.

The ES model was originally built to replace the Military Occupational Specialty Level System (MOSLS) that was built in the early 1970s by General Research Corporation, which is now a part of AT&T Government Solutions [2]. MOSLS was an earlier generation of the current ES model and had essentially the same mission to balance Military Occupation Specialties (MOS) and grade level requirements with the available population of Soldiers. AT&T Government Solutions continues to provide direct support to the Army G1 when they are exercising the model.

Every month the Army G1 uses the Enlisted Specialty (ES) model. The ES model consists of a simulation and optimization that forecasts the Army's enlisted manpower program by MOS and grade across a 7 year planning horizon. The ES model simulates the predicted flow of Army personnel on a monthly basis using historical data to determine the rates and factors for future transactions. Personnel inventory is comprised of two components, the individual account which is made up of Soldiers not available for operational assignments due to training, transition, holdee status or student status, and the operating strength account which is made up of Soldiers available for assignment against an authorization

The optimization portion of the model minimizes the absolute deviation between the operating strength portion of the personnel inventory and the authorizations to best meet the Force Structure requirements while satisfying all the constraints. The objective function in the ES model is to minimize the Operating Strength Deviation (OSD), which is the absolute deviation between the operating strength portion of the personnel inventory and the strength authorizations. Minimizing the OSD is goal of the Army G1 in meeting the Force Structure requirements while satisfying all the constraints. Once the ES model has run to completion, the resulting manpower inventory (by month, skill, and grade) are analyzed and become input for the Analyst Projection Assistance System (APAS) in Human Resources Command (HRC) to be used for personnel distribution planning.

The objective function in the ES model is a weighted sum of the decision variables in the model. The weights of the decision variables are known to change the outcome of the optimization, but it is unclear which weights have the most impact on the resulting OSD. The fundamental questions in MAJ Erdman's thesis are the following:

1. What are the objective function coefficients that have the greatest effect on the absolute deviation between the operating strength and the authorizations?
2. What objective function coefficients are robust with respect to deviations from target strength?

Answers to these questions are expected to help ensure that the target number of Soldiers with the correct skill sets and grade are met. Conducting data analysis necessary to answer the first question is the focus of the work for Team 5 during IDFW 20.

The next section provides a brief overview of the methodology including the experimental designs conducted followed by the results of the data analysis. Finally, insights gained from the workshop and follow-on work are discussed.

METHODOLOGY

The ES model consists of 859,633 variables with 224,473 constraints. Several iterations of the optimizer and simulation are used to converge on a feasible solution. The optimizer prescribes promotions, accessions and reclassifications [2]. The simulator is used to adjust for changes in behavior due to different promotion, accession, and reclassification programs within the Army. The optimization model is solved in CPLEX. A number of iterations of the optimization are performed in order to converge on an optimal solution. The final iteration of the optimizer produces a forecast that is an integer value and resolves any final discrepancies in the ES projections. As the program is currently configured it takes approximately four hours to determine rates and factors and then 17 hours for the model to process through all 15 simulation and optimizations iterations.

In order to meet the objectives of the first research question, traditional experimental design techniques were followed. Design of Experiment (DOX) is a systematic way of exploring a problem where variations are present. The experiments are designed so they can conduct simultaneous examination of multiple factors and explore input factors and their relation to output responses. This allows researchers to identify, compare, and contrast current values while minimizing the number of experiments that need to be conducted. Practicing good experimental design techniques allows for the most cost-effective (in terms of computer processing time, money, etc.) collection of data for future analysis.

Experimental design methodology was used by executing the steps below:

Step 1: Identify input factors, output factor(s) (response variable)

Step 2: Selected ranges that the input factors can take on

Step 3: Identify a screening experiment that will allow the estimation of main effects and potentially two factor interactions

Step 4: Run experiments

Step 5: Analyze data from experiments

Step 6: Based on results suggest an additional experiments required

The input factors are the 52 coefficient values in the objective function, which are presented in Figure 1. The output response is OSD. The levels for each of the 52 input factors are also presented in Figure 1.

A screening experiment allows the researcher to search for a subset of effects that have the most influence on the response variable. The goal of the first research objective in this work is to determine which of the 52 Objective Coefficient Variables were of importance in terms of the response variable, OSD. A Plackett-Burman design was used to study this. The Plackett-Burman is a non-regular factorial design with a low number of experimental requirements, which was important in the case of the ES model because of the long simulation run length. A non-regular design is one that involves partially confounded factors. The Plackett-Burman design created consisted of 56 runs.

Results of the Plackett-Burman design as well as a small set of additional experiments that were conducted during IDFW 20 are presented in the next section.

RESULTS

This section presents the results of the initial experimental design and provides a brief description of the additional experiment and follow-on analysis conducted during the workshop.

Grade Coefficients	Min Range	Max Range	Default Value
Promotion Factors E3	-200	-1	-120
Promotion Factors E4	-30	-1	-4
Promotion Factors E5	-30	-1	-4
Promotion Factors E6	-30	-1	-4
Promotion Factors E7	-30	-1	-4
Promotion Factors E8	-30	-1	-4
Promotion Factors E9	-30	-1	-4
Reclass Factors E3	3000	7000	5000
Reclass Factors E4	1	400	5000
Reclass Factors E5	1	400	200
Reclass Factors E6	1	400	200
Reclass Factors E7	1	400	200
Reclass Factors E8	1	400	200
Reclass Factors E9	1	400	200
Strength Factors E3	0	200	0
Strength Factors E4	1	4000	2000
Strength Factors E5	1000	5000	3000
Strength Factors E6	2000	6000	4000
Strength Factors E7	3000	7000	5000
Strength Factors E8	4000	8000	6000
Strength Factors E9	5000	9000	7000
Target Factors E3	0.01	1	0.1
Target Factors E4	0.01	1	0.1
Target Factors E5	0.01	1	1
Target Factors E6	0.01	1	1
Target Factors E7	0.01	1	1
Target Factors E8	0.01	1	1
Target Factors E9	0.01	1	1

Non-Network Constraint Coefficients	Min Range	Max Range	Default Value
MOSS Shred Target	0.01	0.32	0.13
NCO Constraint	0	0	0
SL1 Constraint	0.1	0.5	0.3

Objective Function Arc Coefficients	Min Range	Max Range	Default Value
Conversion	-9500	-5500	-7500
MOS Deletion	0	0	0
Demotions	-6000	-6000	-6000
LOSSES (ETS)	-9999	-8000	-9999
LOSSES (Other)	-9999	-8000	-9999
LOSSES (PML)	-50	-50	-50
LOSSES (Retirement)	-9999	-8000	-9999
LOSSES (Training)	-9999	-8000	-9999
Promotion (VBU)	-7000	-3000	-5000
Percent Bound on Pool Flows	100	100	100
Reclassification (Mandatory)	-9999	-8000	-9999
Reclassification (Reenlistee)	-9999	-8000	-9999
Reclassification (Voluntary)	0	100	1
Same-Grade Reclassifications	0	100	10
THS	-9100	-5100	-7100
BCT Training	0	65	31
NPS Without Training	1200	5200	3200
OSUT Training	0	65	33
PS with Training	0	65	34
PS Without Training	1500	5500	3500
Training Deletions to Pools	0	80	40

Figure 1: 52 Objective Coefficient Variables with minimum, maximum, and default values

A graphical representation of the OSD response for each of the 56 runs from the initial Plackett-Burman design are depicted in Figure 2.

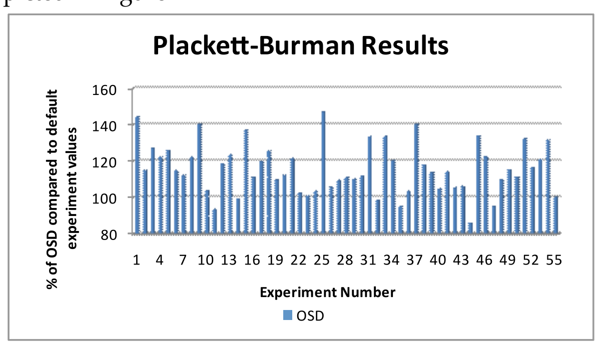


Figure 2: Placket-Burman Results

The software package JMP®, a product of SAS Institute, was used to analyze the data from the experimental design runs. Initial analysis consisted of performing a stepwise regression. The inputs to the stepwise regression included all of the main effects and two factor interactions. Note that this amount of terms indicates that the design is super saturated. The stepwise linear regression in JMP uses lengths method to identify statistically important coefficients.

Once the Stepwise regression results were completed Least Squares regression was used to build a linear regression model including only the significant terms as indicated by the stepwise regression results. The significant input factors are presented in Figure 3.

Parameter Estimates

Term	Estimate
Intercept	352091.21
Strength Factors E4	3428.0933
SL1 Constraint	-1130.819
Promotion Factors E5*Promotion Factors E6	-2845.178
Promotion Factors E8*Target Factors E4	5925.8127
Reclass Factors E5*Strength Factors E5	2284.1627
Reclass Factors E6*Target Factors E7	-2795.389
Reclass Factors E6*SL1 Constraint	1341.4701
Reclass Factors E9*Strength Factors E8	-2333.568
Strength Factors E3*SL1 Constraint	-3604.157
Strength Factors E4*LOSSES (ETS)	-2307.061
Strength Factors E4*NPS Without Training	-2806.16
Strength Factors E5*PS Without Training	-2011.372
Strength Factors E7*Target Factors E9	17766.789
Strength Factors E8*PS Without Training	4506.6355
Demotions*Reclassification (Reenlistee)	1662.5679
LOSSES (Retirement)*NPS Without Training	1189.8346
BCT Training*OSUT Training	9892.3711

Figure 3: JMP® Output of Significant Factors from Plackett-Burman Experiments

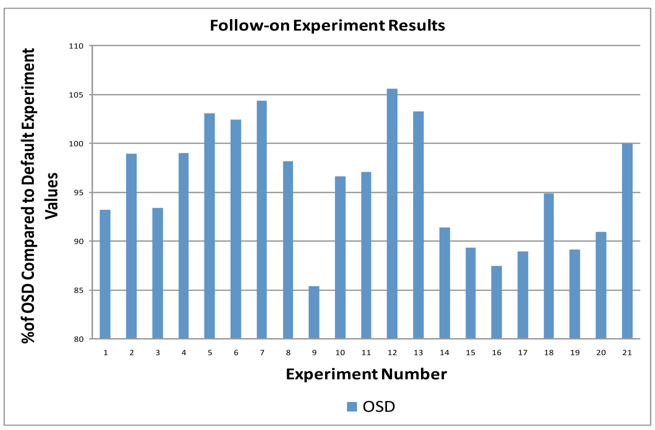


Figure 4: Follow-on Experiment Results

In order to ensure that these terms are in fact significant, follow-on experiments were conducted. The follow-on experiment consisted of changing only the top nine factors from the previous experiment and holding the other 42 coefficients at their default values. A space filling experiment was used in order to provide more degrees of freedom to test for significance of higher order polynomial terms and to provide any de-aliasing necessary for the terms identified as significant. Once complete the results of the 20 runs were processed and compared to the default coefficient OSD listed as experiment 21 on Figure 4.

Manipulation of the nine coefficients in the follow-on experimental design resulted in 75% of the OSDs being below the current default OSD. These results are encouraging and show that these nine coefficients are important and can be used to reduce the overall OSD in future experiments.

All of the experiments completed were cross validated to see which coefficients (linear, squared, or interaction) were robust with respect to predictive abilities. The data points were placed into JMP® except for 10 randomly excluded points. Stepwise regression was executed and JMP® selected the coefficients and that played a significant role in predicting the OSD. To prevent over fitting only the top 10 significant terms were taken from the Stepwise regression and used in the Least Squares regression. Limiting Least Squares to only the top 10 coefficients eliminated the problem of over fitting the data but still resulted in the R2 and adjusted R2 being above the .90 level.

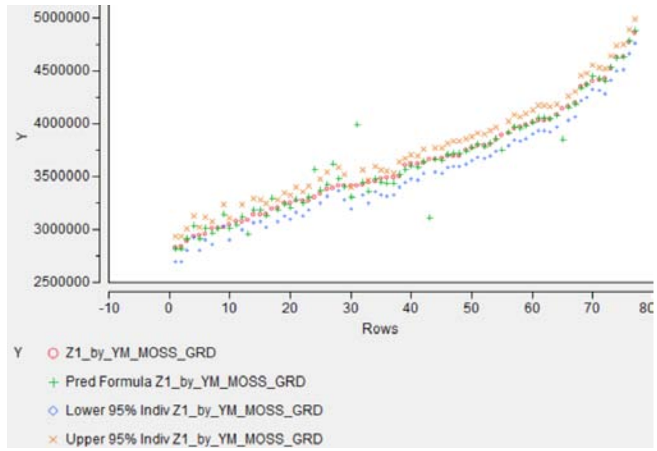


Figure 5: Cross validation plot

CONCLUSIONS

The DOX principles guided the execution of experiments on the ES model and ensured a comprehensive exploration of the problem space and efficient use of computer processing resources. The DOX provided valuable insight into how the coefficient inputs affect the OSD. The initial screening experiments also highlighted what areas require additional experiments.

Based off the work conducted at the IDFW, Team 5 was able to illustrate that the ES model outcome can be predicted by using a small subset of the significant coefficients. The cross validation of the current work shows that the coefficients still require more experimentation in order to produce a good working model for predicting the OSD. Based off the work conducted here additional experiments will be executed and a working mathematical model will be formulated.

REMARKS

The research into the ES model is ongoing and is expected to be completed by the end of June. The hope continues to be that this research will gain new insights into the ES model and help the United States Army personnel optimization.

REFERENCES

[1] http://www.armyG1.army.mil/

[2] Hall, Andrew O. Validation of the Enlisted Grade Model Gradebreaks. Winter Simulations Conference, 2004: 921-925

[3] Montgomery, Douglas C. Design and Analysis of Experiments 7th Edition. New York: John Wiley & Sons, 2009.

Team 6: Applying Social Network Analysis to Data Farming of Agent-Based Models

TEAM 6 MEMBERS

Meyer, Ted
McDonald, Mary
Upton, Steve
Naval Postgraduate School, US

Middleton, Donna
Northrop Grumman, US

Thomas Holland
Naval Surface Warfare Center, Dahlgren, VA

Kim Lewis
Department of Defense

Harald Schaub
IABG, Germany

Fernando Freire
National Defense College, Portugal

Christina Bouwens
SAIC, US

INTRODUCTION

Team 6 continues to participate in an ongoing study to examine the utility of distillation modeling in the Counter-IED (Improvised Explosive Devices) fight. Understanding social networks, their nature in insurgencies and IED networks, and how to impact them, is important to the Counter-IED battle. Team 6 is exploring methods of extracting, analyzing, and visualizing dynamic social networks that are inherent in agent-based models in order to build tools to examine and manipulate insurgencies. We are starting with basic clique creation scenarios as the initial basis of our investigations and are examining the types of network statistics that can be used as MOEs and pointers to unique and emergent behaviors of interest.

The Team 6 goals during IDFW 20 were to extend our base scenario with simple variations and to test candidate tools and prototype methods for data farming the scenario, extracting network data, analyzing end-of-run network statistics, and visualizing network behaviors.

Social Network Analysis (SNA) techniques were explored in detail to determine which network metrics would be most beneficial for analyzing the types of networks produced by our agent based model. This would allow the team to explore questions regarding Counter-IED issues—including insurgent network evolution and adaptation. Within insurgent, IED-using networks, there are two of interest: IED Emplacement Networks (consisting of personnel that are directly involved with IED usage) and IED Enabling Networks (consisting of communities that indirectly support the IED Emplacement networks). Team 6 is in the process of identifying tools that can be used to explore patterns that might provide valuable insights into emergent behaviors of interest.

Background

In previous work, at IDFW 19 and between workshops, the team:

- Examined a set of agent-based model (ABM) C-IED (Counter-IED) task plans generated by previous workshops.
- Selected potential candidate tasks for follow–up study and analysis.
- Concluded that SNA concepts and techniques needed to be applied to address the candidate tasks.
- Demonstrated the ability to extract social network data from a basic social interaction scenario.
- Data farmed initial scenario and established need to simplify the target scenario in order to more closely examine cause and effect relationships to SNA statistics.
- Developed a new base scenario, delineated a simple illustrative DOE, and data farmed the model to provide a sample data set for further exploration.

IDFW 20 Objectives

Team 6's objectives for IDFW 20 were to:

- Examine utility and approach of applying specific SNA statistics, methods, and concepts using the data farming output provided from previous work.
- Delineate the data requirements for the various types of networks that might be extracted from modeling.
- Establish and document software and processes for applying these capabilities to detecting and analyzing emergence.

To address these objectives, the team started with a very basic approach. Assuming that an agent based simulation produces a time-series of state data and MOEs, our tools and methods need to allow the analyst to conduct tests to:

- Detect the presence of a network or networks.
- Distinguish different networks and different classes of networks.
- Determine if and when networks achieve equilibrium.
- Determine which model inputs have significant impact on the state and behaviors of the network.

Specifically, the intent is to use these capabilities to be able to address a variety of social network questions such as:

- What do insurgent networks look like? Who is in the network? Who is not?
- How do we distinguish networks that should be attacked, networks that should be attritted or that should be co-opted?
- Who are the High Value Individuals (HVI) and what are their identifiable characteristics?
- Will removing specific nodes destabilize a network?
- What are the 2nd and 3rd order effects?
- What are the potential unintended consequences?

Abstracted Illustrative Scenario and DOE

Initial work was based on the Pythagoras distribution "Peace" scenario. Data Farming of this scenario and initial analysis of the results between IDFW 19 and 20 led to the development of a more basic scenario in order to test basic network concepts.

The illustrative "Clique Creator" (CC) scenario was developed using Pythagoras's "relative" color change capability as a tool for experimenting with SNA extraction and analysis. CC has a single agent class with 100 instantiated agents that are uniformly distributed across Pythagoras's red and blue color spaces. The agents' only "weapon" is "Chat" which induces a relative color change on other agents with which the agent interacts. As the scenario is executed, entities move through various color states, becoming "more" red or "more" blue depending on the interactions with other red or blue entities. States will change depending on whether two entities engage in "chatting" and form a connection. The more any two agents interact, the more "alike" they become.

The focus of the scenario selection was to represent dynamic homophily and use the results to explore the various analysis tools under study. Multiple excursions / replications of the Pythagoras-developed Clique Creator scenario were used to produce the data for analysis with the candidate tools. This baseline provided a means for the team to experiment with various SNA measures and analysis techniques.

Pythagoras can provide multiple views of agent state data. A spatial view showed the physical relationship between entitities and where connections or bonds were formed. The inclination space view sorted the entities by colors. This color space view is used to illustrate the homophilic state of the participating entities in the simulation.

A very basic full-factorial design space was used to data farm the scenario. The design matrix (Table 1) reflects four input parameters that will influence the composition of the resulting networks:

- RelativeChange - Percentage relative change of color when "chatted."
- InfluenceRng - Maximum distance of chat.
- FriendThresh - Agents within this range are considered "linked."
- EnemyThresh – Dependent variable; is calculated as FriendThresh plus 55, in order to preserve the same Friend to Enemy Distance (equivalent to the "neutral" range) as was present in the base scenario.

Excursion	RelativeChange	InfluenceRng	FriendThresh	EnemyThresh
0	5	25	5	60
1	5	25	50	105
2	5	25	100	155
3	5	100	5	60
4	5	100	50	105
5	5	100	100	155
6	5	250	5	60
7	5	250	50	105
8	5	250	100	155
9	20	25	5	60
10	20	25	50	105
11	20	25	100	155
12	20	100	5	60
13	20	100	50	105
14	20	100	100	155
15	20	250	5	60
16	20	250	50	105
17	20	250	100	155
18	60	25	5	60
19	60	25	50	105
20	60	25	100	155
21	60	100	5	60
22	60	100	50	105
23	60	100	100	155
24	60	250	5	60
25	60	250	50	105
26	60	250	100	155

Table 1 – Clique Creator Experimental Design Matrix

The CC scenario can be considered as a metaphor for a group of people establishing relationships based on shared interests or desires (color space proximity) and physical proximity (relative agent location). Agents are drawn toward agents with similar color and move away from agents of disimilar color. The closer agents are in location, the more frequently they "chat" each other, and thus, the closer they grow in color space. Eventually, cliques of "like-interest" agent form and are impacted by other agents and cliques. The input parameters varied in the design matrix affect these behavioral processes in straightforward ways.

Visualizing the Dynamic Network State

Part of a toolset to examine social network dynamics is the ability to analyze the ongoing agent interactions, behaviors, and network responses. Co-visualizing the various aspects (layers) of network dynamics can potentially provide powerful insight into the network. Team 6 has done initial examination of the CC scenario using several visualization capabilites. Figure 1 is the spatial view provided by Pythagoras.

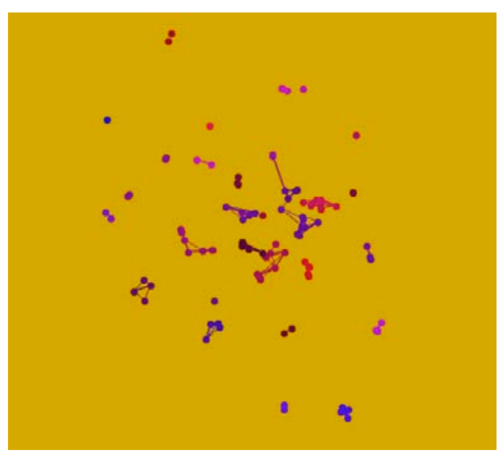


Figure 1 – Clique Creator Scenario – Spatial View

Figure 1 shows the agents at a time-step midway in the scneario. "Chats" are shown as lines between agents. This view, though, focuses on the location of the agent spatially.

Figure 2 shows four time-steps of an "inclination"-space view. In this image the location of the agents is based on their location in color space. The "redness" (0-255) of the agent is represented on the x axis. The "blueness" (0-255) of the agent is represented on the y axis. As the scenario proceeds left to right, top to bottom, note the congregation of agents into color groups. These groups do not represent the cliques formed though, because the spatial aspect is not represented.

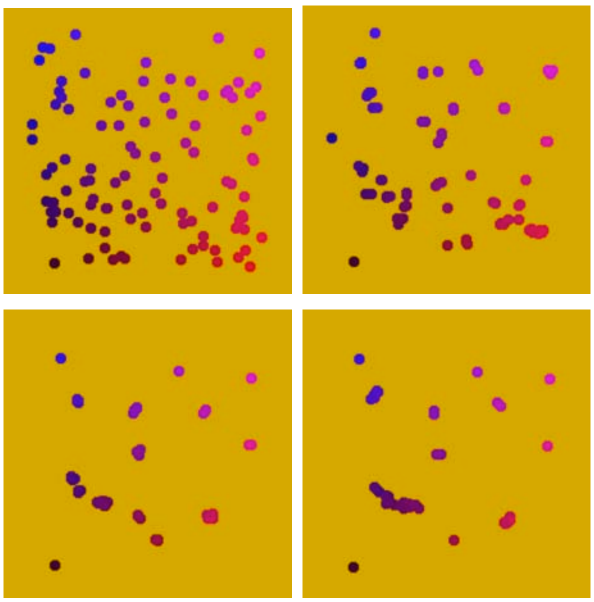


Figure 2 – Clique Creator Scenario – Inclination Space View

FIgures 3 and 4 represent the same agent network , derived from the CC scenario, using the social network analysis "layout" generated by the R SNA plug-in (http://cran.r-project.org/web/packages/sna/index.html) and SoNIA (http://www.stanford.edu/group/sonia/index.html) software packages.

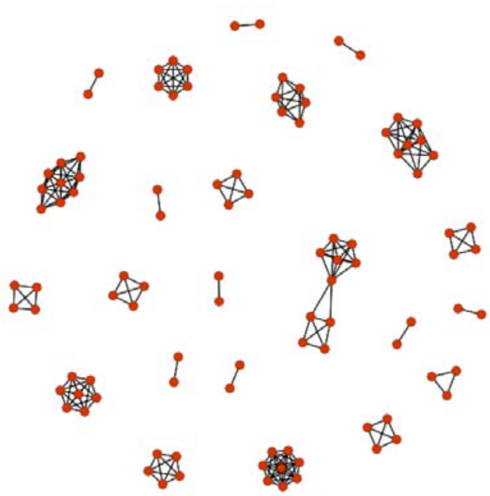


Figure 3 – Clique Creator Scenario – Static Graph View

Figure 3 shows a static network layout representation of one of the CC time-steps using the default SNA layout algorithm. The SNA R package plots each time-step independently, not accounting for the layout defined in the previous time-step. As a result, the dynamic evolution is difficult to examine.

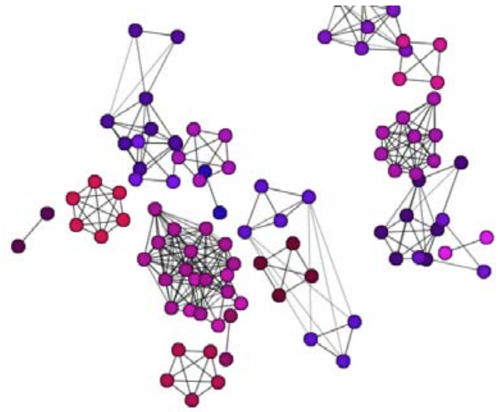


Figure 4 – Clique Creator Scenario – Dynamic Graph View

Figure 4 shows a single time-step using the SoNIA application. SoNIA is designed to support dynamic time-series network data. As a result, the layout of any timestep is based on the previous time step as a starting point. The result is a layout which displays the evolution of the network, but that can result in layouts that are not easily viewed statically.

It should be noted that Figures 2, 3 and 4 do not represent the spatial data shown in Figure 1 in any way... the "physical" location is ignored in these representations. In Figure 2 location represents color, and in Figures 3 and 4 the location is purely a function of the layout algorithm, which is designed to display the network in an uncluttered and easily-viewed manner, not the spatial location of the agents.

Social Network Analysis (SNA)

One of Team 6's goals is to begin to understand the utility of various SNA statistics in understanding the scenario dynamics and the result of data farming. Step one in this process during this workshop was to delineate what outputs and analysis methods provide insight into network evolution and impact on agent behaviors.

SNA statistics fall into two classes: node statistics and network statistics. Node statistics include: betweenness, closeness, eigenvector centrality, and degree. Network statistics include: number of components, number of cliques, and average path length.

The team decided to focus on node statistics initially and produced time-series output for every node of betweenness, eigenvector centrality and degree. Although data for 27 excursions of data farming was collected, it was decided to do an intial comparison of three excursions, where the primary variation was the color distance that defined what is considered a friend (a homophilic link) . Excursions 1, 2, and 3 were examined.

Figures 5a, 5b, and 5c represent a single replication of excusions 0, 1, and 2 as delineated in Table 1. The plots represent the degree of each agent over time. The vertical axis is degree (the number of links associated with a node), the horizontal axis is time, and the axis going into the page is agent number. Figure 5 was generated using the PlotGL plugin to R.

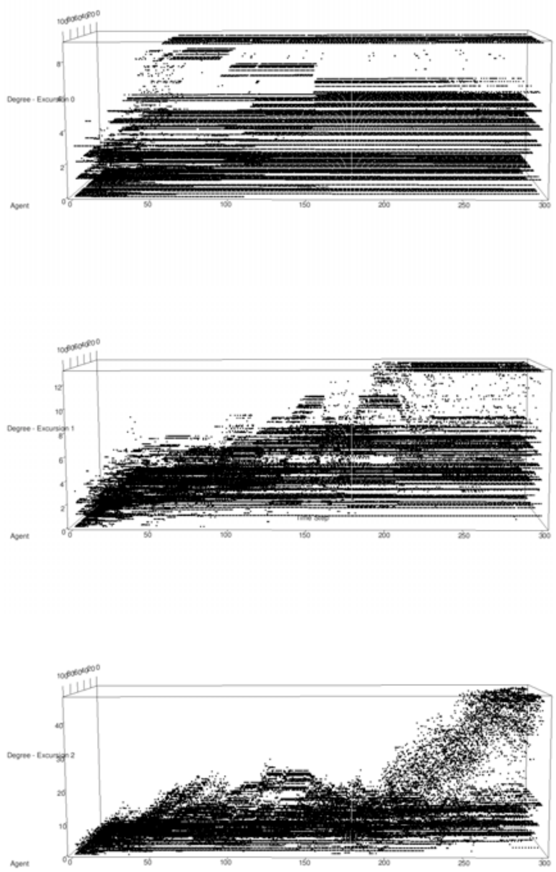


Figure 5 – Degree Centrality for Excursions 1-3

In Figure 5, various pattern differences, related to the evolution and devolution of cliques and components, can be discerned There are obvious differences between the excusions, with 0 and 1 appearing to reach covergence, but 2 never converging. It can be seen that some agents reach a steady-state and maintain it for some time, while other groups of agents particpate in behaviours which lead to the growth and reduction of degree for groups of agents.

Surprises

Two surprises (counter-intuitive results) presented themselves. Excursion 2, in Figure 5c, shows that an increase in FriendThresh, that is, expanding the range and number of agents that an agent has homophilic links with in color space leads to increased instability in terms of clique formation. The initial assumption was that this would affect the size of the cliques and number of components. The unexpected result is that this increase prevents the stabilization of cliques and network components. Rather, it appears that this increase results in groups being able to "steal" members from other groups more easily.

Another interesting behavior is the Excursion 0 (Figure 5a) degree variation that occurs before equilibrium. In this case it appears that larger components are formed intially, but that they devolve into smaller groups over time. The team intends to investigate the set of replicates associated with this excursion to determine whether this behavior is consistent for this level of FriendThresh.

Summary and Way Ahead

Significant insight was gained by team members in delineating capabilities needed in a toolkit for the extraction and analysis of dynamic social data from models. The following capabilities will be needed for ongoing data farming research of basic social networks:

- Synching of Visualization: Various representations of the dynamic network are useful, but examining multiple views of the network time-step synced would provide powerful relational insights.
- Equilibrium Time: Determining whether equilibrium occurs and how long it takes is often the first step in analysis.
- Data Farm Time Window Reduction Size: Dynamic network analysis requires defining what constitutes a link, for example, a single interaction or multiple interactions over some time window. Being able to data farm this time window would provide analysts insight into network basics.
- Node Statistic Capability: Degree, betweenness, eigenvector, closeness need to be extractable for each node, time-step, replicate and excursion and then represented effectively.
- Network/Component Statistic Capability: # cliques, and components, density, and others need to be acquired for each time step, replicate and excursion.
- Newcomer/Leaving Effects: Measure the effects of dynamic birth and death of agents.
- Network Boundary Effects: Data farm the impact of varying the size and extent of the network.
- MOEs (end-of-run vs. time-series).

Team 6 will continue to delineate tool capabilities for data farming social network models. We intend to accomplish the folowing tasks in the upcoming months:

- Document tools and methods identified in IDFW20.
- Define model output requirements for SNA analysis.
- Expand toolkit to include additional network, node, and link statistics.
- Expand data farming methods for other network layers including weapon and resource interaction, spatial, communication, and multiple "inclination" parameters.
- Continue detailed analysis of CliqueCreator data farming results.
- Test use of tools and methods on other models (MANA, Netlogo scenarios).
- Begin delineating insurgent IED network scenario.

Team 7: Data Farming to Support Model Validation of the BTRA-BC Battle Engine (BBE)

TEAM 7 MEMBERS

Blais, Curtis
Stork, Kirk
Upton, Steve
Naval Postgraduate School, US

Eaton, Jeffrey
Hoffman, Matthew
Concurrent Technologies Corporation, US

Rollins, Steven
USMC Logistics Command, US

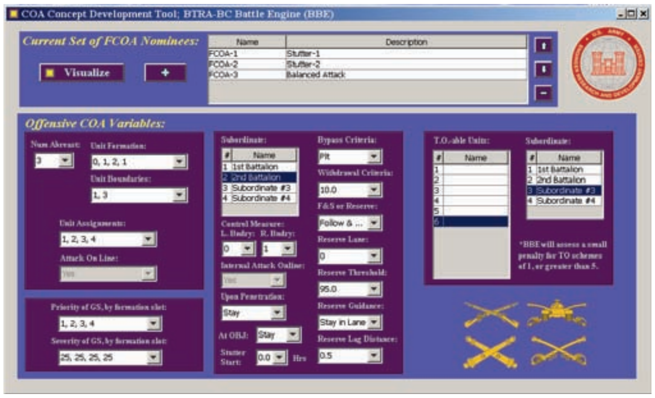


Figure 1: BBE Friendly COA definition window

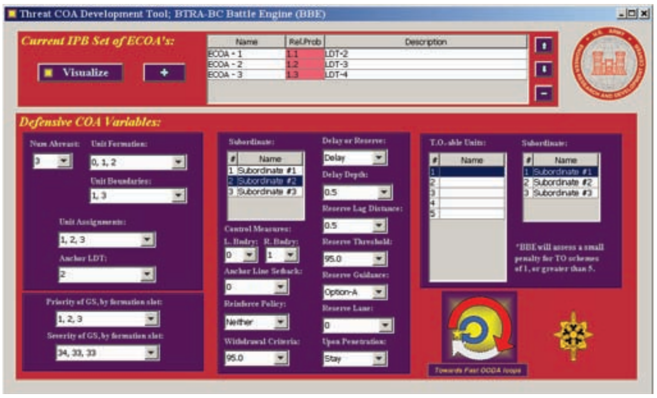


Figure 2: BBE Enemy COA definition window

INTRODUCTION

The Battlespace Terrain Reasoning and Awareness Battle Command (BTRA-BC) Battle Engine (BBE) [1] is a software tool designed to assist commanders and staffs in developing and analyzing Friendly Courses of Action (FCOAs) in the context of mid-to-high intensity combat operations.[1] It is designed to automate a number of sub-tasks of the Military Decision Making Process (MDMP) [2] that previously have been the exclusive domain of the human planner. Using BBE, commanders and staffs can quickly generate and evaluate an unprecedented number of FCOAs. BBE is intended to increase the speed of tactical decision making without sacrificing the quality of those previously manually-developed alternatives.

A major subcomponent of the MDMP is the Intelligence Preparation of the Battlefield (IPB) process [3], culminating in development of Enemy Courses of Action (ECOAs). This process mirrors FCOA generation, but is focused on identification and evaluation of potential enemy activities. A simplified set of procedures and analysis tools within BBE can also be used for generating ECOAs.

Gaming, in the most basic sense, is an attempt by one player to devise and implement a strategy to defeat an opponent or overcome a set of circumstances. In its most basic form, a game requires game pieces, a game environment or game board, and game rules. BBE game pieces represent military units, both enemy and friendly, and interactions between these pieces represent the fire and maneuver of tactical combat operations. The BBE game board represents an abstraction of the battlespace that preserves those tactical aspects of terrain that influence tactical operations. The game board greatly resembles the traditional Modified Combined Obstacle Overlay (MCOO) produced during the IPB process. The BBE reference model governs how game pieces interact with each other and with the game board; it represents the "rules" of a BBE game run.

Objectives

Team 7 sought to achieve the following objectives:

- Use design of experiments and data farming to support BBE validation
- Key Question: What are the factors having the greatest effects on BBE-computed outcomes and the scoring of FCOAs against specified ECOAs?

[1] Our thanks to the developers of the BBE tool, Mr. Jerry Schlabach and Mr. Eric Nielson, from the U.S. Army Geospatial Center, for support in learning the operation of the tool, creation of a version of the software for use in conducting runs of the software separate from the Graphical User Interface, and for the description of the software used in this paper's Introduction.

Model Validation

The team members were drawn to Team 7 by an intense interest in the use of data farming as a tool in model validation. On the one hand, Mr. Blais, Mr. Stork, and Mr. Upton are members of an NPS team that have been funded by the Army Geospatial Center to perform various validation studies on the BBE tool to develop evidence that can be used by accreditation authorities to determine if the tool is fit for its intended purpose. Mr. Eaton, Mr. Hoffman, and Mr. Rollins came from a background of common work on validation of models for the USMC Logistics Command (LOGCOM). Discussions in early working sessions of the team dealt with the concept and practice of model validation (in some cases, in contrast to the concept and practice of model verification). In [4], Dr. Petty describes two principal comparisons as the focus of validation activities: (1) comparison of the real world to the conceptual model; and (2) comparison of results from the executable model to the real world. This description is refined in [5] to (1) comparison of the referent(s) (i.e., what is known about the real world relative to the intended use of the model) to the conceptual model; and (2) comparison of the results from the executable model to the referent(s). In light of these considerations, the team wanted to generate evidence that would support a decision that the model is (or is not) useful for its intended purpose through the following actions:

- Investigate computational behavior (sensitivities) of the model for expected and anomalous outcomes
- Confirm expectations of the BBE conceptual model
- Provide a "conversation-starter" between the software developer and the validation team to advance common understanding of the intended model behavior.

Our purpose was definitely not to conduct verification studies with the model; that is, we were not investigating the correctness of the implementation of the software logic with respect to the conceptual model. On the other hand, it was recognized that any "disconnect" found by data farming, in light of expectations raised by knowledge of the conceptual model for the tool, either could be indicative of an issue in the implementation itself (a possible verification finding) or indicative of an issue in the conceptual model (a possible validation finding). In either case, the objective of the study with respect to model validation was not to judge whether the outcomes were right or wrong (against some criteria), or good or bad (based on some valuation), but to produce evidence that could be used by others in position to make such assessments with respect to the intended use of the tool.

Data Farming Approach

All but one of the members of Team 7 were "IDFW Rookies;" the one exception being Steve Upton from the NPS SEED Center. This meant that the team was particularly motivated to internalize the guidance and best practices of data farming as described in one of the plenary sessions and demonstrated throughout the working sessions by Steve. Perhaps the following summary will be useful to readers of this article, and potential attendees of future IDFWs:

- Determine an initial set of factors to explore. Because the foundation of the BBE processing logic is the underlying representation of important features of the terrain (maneuver network), we decided to begin by examining the sensitivities of model outcomes to

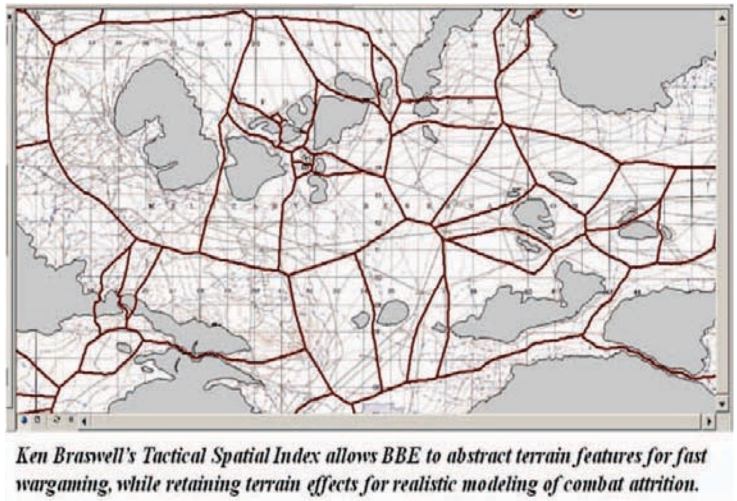


Figure 3. Terrain Analysis Mobility Network

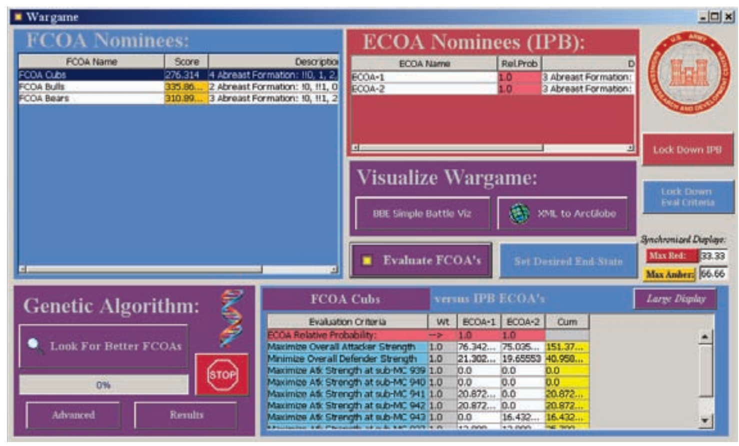
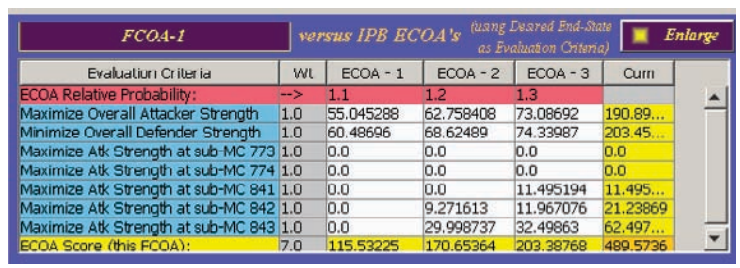


Figure 4. Wargaming FCOAs against ECOAs in BBE



Evaluation Criteria	Wt	ECOA - 1	ECOA - 2	ECOA - 3	Cum
ECOA Relative Probability:	-->	1.1	1.2	1.3	
Maximize Overall Attacker Strength	1.0	55.045288	62.758408	73.08692	190.89...
Minimize Overall Defender Strength	1.0	60.48696	68.62489	74.33987	203.45...
Maximize Atk Strength at sub-MC 773	1.0	0.0	0.0	0.0	0.0
Maximize Atk Strength at sub-MC 774	1.0	0.0	0.0	0.0	0.0
Maximize Atk Strength at sub-MC 841	1.0	0.0	0.0	11.495194	11.495...
Maximize Atk Strength at sub-MC 842	1.0	0.0	9.271613	11.967076	21.23869
Maximize Atk Strength at sub-MC 843	1.0	0.0	29.998737	32.49863	62.497...
ECOA Score (this FCOA):	7.0	115.53225	170.65364	203.38768	489.5736

Figure 5. Friendly COA Evaluation Scores in BBE

settings of the five terrain modifiers implemented in the model.

- Use the SEED Center's Nearly Orthogonal Latin Hypercube spreadsheet tool [6] to generate design points for the five terrain modifiers characterizing the maneuver network. We generated 513 design points across a range of values from 0.1 to 2.0 for each modifier.
- Use the SEED Center's OldMcData tool [7] to generate XML (Extensible Markup Language) excursion files from a base BBE scenario definition file (also in XML).
- Execute the model once (since the model computation is deterministic) for each of the excursion files.
- Use a post-processing tool (customized script developed by Steve Upton) to gather the output data into a single file to load into JMP (http://www.jmp.com) for statistical analysis.
- Perform quantitative and qualitative analysis on the results.
- From examination of various views and analyses in JMP, form new hypotheses possibly identifying other variables of interest. Then iterate the process.

RESULTS AND ANALYSIS

In the span of the workshop's four days, Team 7 was able to iterate over the above process three times, generating approximately 2000 individual outcomes from the BBE model. All of these runs were conducted in the context of a single demonstration scenario provided by the developers of the model. It should be noted that the model runs very quickly; although we used a small cluster for our runs, we could have easily performed all the runs on a laptop. Due to space limitations, we will only discuss our first two iterations in this paper.

By "scenario" we mean a single order of battle, along with a fixed set of three enemy courses of action, and a fixed set of evaluation criteria all taking place in a single physical setting. Initially, the scenario included two human-crafted FCOAs. The FCOAs are identified as "Devin Hester" and "Jay Cutler," or by their abbreviated names ("Devin" or "Hester" for the first, and "Cutler" for the second). In the final iteration of our process we also examined several machine-generated FCOAs. With respect only to the FCOAs and ECOAs, our experimental design is full-factorial.

Initial Line of Inquiry

The variables included in a course of action are many, and the dimensionality is not constant. That is, there are COA choices that add to the number of COA variables in a hierarchical way. Because our team was generally new to Data Farming, we chose to avoid that complexity and focus on a key premise of the model; i.e., that terrain properties are a key contributor to combat outcomes.

The model abstracts terrain into Maneuver Corridors (MC), which represent possible paths over which military units can travel. The MCs have a set of properties that are used by the model to compute rates of advance, and to modify the outcomes of combat activity that takes place within them. The modifiers available for each MC are:

- Road Speed
- Attack Maneuver
- Defense Maneuver
- Attack Fire Support
- Defense Fire Support

The Maneuver modifiers are applied to maneuver units such as tanks and infantry, and the Fire Support modifiers are applied to fire support units, such as artillery.

Our first iteration, then, farmed over these MC multipliers as a way to explore model behavior over various assumptions about the impact of terrain on model outcomes. As mentioned previously, we applied the SEED Center's NOLH DOE tool to explore values between 0.1 and 2.0 on each MC multiplier. This range covers terrain input values that might never occur in practice; however, we wanted initially to explore a full range of possibilities, with the expectation that future iterations will narrow our focus based both on our findings and the advice of subject matter experts.

The outcome of each model run is a score for each FCOA/ECOA pair. These pair-wise scores are also aggregated in a user-defined way that reflects the IPB estimates of the likelihood of encountering each ECOA (i.e., a weighted average). For our runs, we weighted the ECOAs as provided for in our example scenario, so the "overall scores" are based on the weightings in Figure 6.


ECOA Nominees (IPB):

ECOA Name	Rel.Prob	D
Fwd Defense – 3 Abreast	1.0	3 Abreast Formation:
Company sized reserve	2.0	3 Abreast Formation:
Battalion (+) reserve	1.5	3 Abreast Formation:

Figure 6. ECOA Weightings

The first thing we noticed from these runs was the relative unimportance of the Fire Support modifiers. In Figure 7, we show the per-ECOA scores across all the MC Multipliers, where you can see that Road and Maneuver multipliers are the only ones with visible effect. We took particular interest in a second feature, that there seem to be values of the Road Speed modifier that cancel out all other multipliers and clip the score to 250.

Because most of the team was also new to using the JMP software, we devoted a substantial amount of time to exploring the data with this new tool. As we thought about the use of this model as a decision aid, it occurred to us that the score might be of less interest than the relative ranking of FCOAs. With some JMP magic by Steve Upton, we produced Figure 8, which shows a clear break in preference from the Devin FCOA to the Cutler FCOA when the Road Speed multiplier gets below about 0.7.

We were initially somewhat surprised that changes to the Road Speed multiplier would cause such a sharp delineation of the recommended COA. However, upon reflection, this result is not so surprising. The model is telling us that not only does the Devin COA capitalize on high mobility terrain,

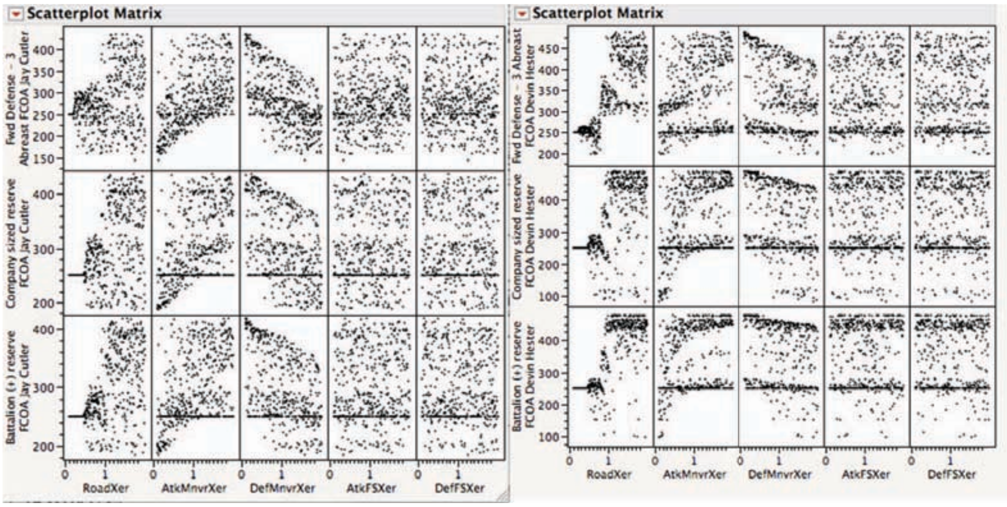


Figure 7. Scatterplots of FCOA Scores versus Terrain Factors

New Questions for Iteration Two

The BBE model gives the planner a choice for the time resolution of the combat model. Time slices available are 6, 12, 18, 24 and 30 minutes. The importance of the road speed multiplier in the first iteration caused us to wonder how the time slice selected would impact the results. We formulated two new research questions.

a. Is the outcome of the model (i.e., rank and/or scoring of a COA) dependent on the time slice selected?

b. Are the conclusions about why one COA is better than another consistent across all the time slices? (i.e., is the conclusion about Devin's superiority in high road speed multiplier consistent as we change the time slice?)

but that it may be a very bad choice when mobility is constrained. This conclusion of the model can now be subjected to expert criticism; for example, all things considered, for this specific scenario, do experienced tacticians agree with this conclusion?

As newcomers to Data Farming, we found this outcome to be encouraging with respect to using data farming as a tool to support questions of model validation.

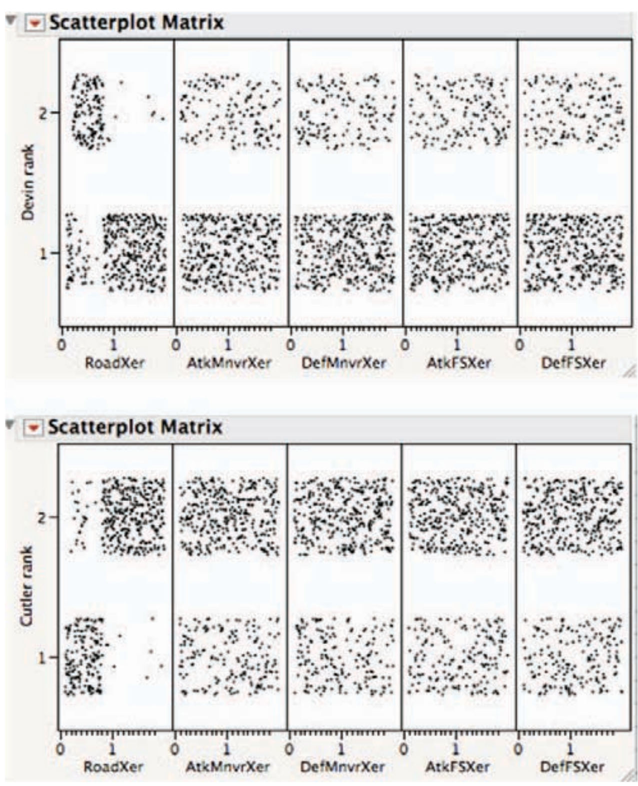


Figure 8. Scatterplots of FCOA Rank versus Terrain Factors

To perform this test, we examined the outcomes across the 5 time slice options provided in the model.

Figure 9 shows the score distribution (aggregated scores, recall Figure 6) for two time slice selections. Clearly Devin does better in the upper frame, which is a time slice value of 12 minutes, than it does in the lower frame, where the time slice is 6 minutes.

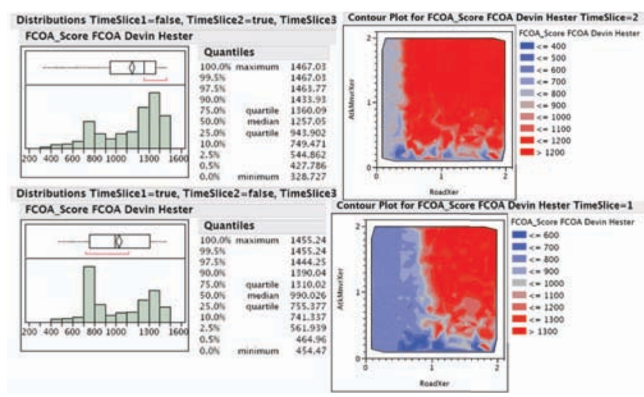


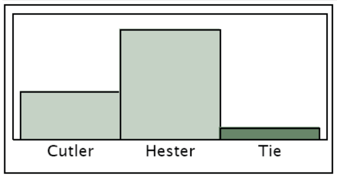
Figure 9. Distribution of FCOA Scores by Timeslice Value

We again return to the purpose of the model as a decision aid and investigate the relative rankings. Figure 10 confirms that Hester (i.e., Devin) does relatively better with a larger timeslice, winning 443 times in a 12 minute time slice and only 339 in a 6 minute time slice.

In retrospect, we would not use the aggregated scores for this analysis. As we are attempting to validate the underlying combat resolution mechanism, dealing with the probabilities of encountering any particular ECOA only serves to cloud the results. If we can make sound statements about the model's recommendations for each FCOA/ECOA pair, then we have done our job. The validity of weighting by relative probability of occurrence is a separate matter.

Distributions TimeSlice=1

Overall Rank

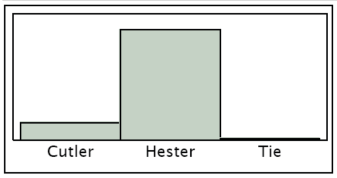


Frequencies

Level	Count	Prob
Cutler	144	0.28070
Hester	339	0.66082
Tie	30	0.05848
Total	513	1.00000
N Missing	0	
3 Levels		

Distributions TimeSlice=2

Overall Rank



Frequencies

Level	Count	Prob
Cutler	68	0.13255
Hester	443	0.86355
Tie	2	0.00390
Total	513	1.00000
N Missing	0	
3 Levels		

Figure 10. Distribution of FCOA Ranks by Timeslice Value

Perhaps of even more interest is the second question. This interest is because our intuition would suggest that the implementation detail of selecting a time slice should not cause the model to give different conclusions about how the world works.

T

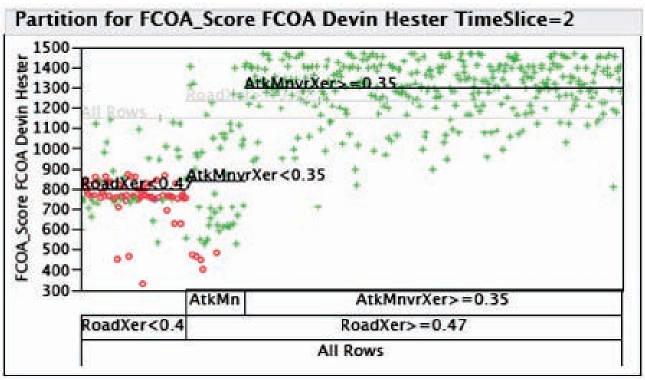


Figure 11. Regression Tree for Selected FCOA and Timeslice 2 (12 minutes)

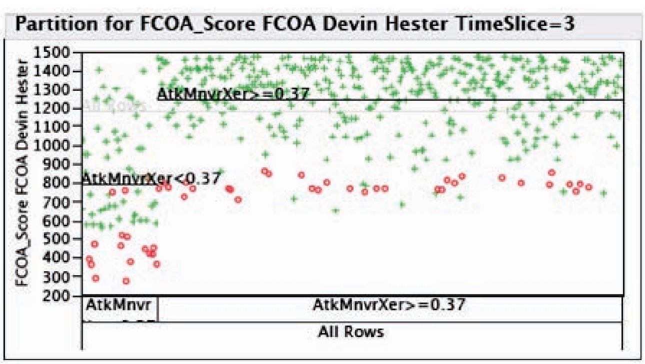


Figure 12. Regression Tree for Selected FCOA and Timeslice 3 (18 minutes)

The results actually show that the important factors in the model change as a result of changing the time slice selection. Figure 11 is a regression tree with time slice set to 12 minutes. This analysis applies equally to the runs performed in our first iteration, and shows the same result. The green FCOA (Devin Hester) performs best when the Road Speed multiplier is greater than 0.4. We remain suspicious that something else may be going on at these low values, since both COAs cluster tightly around the aggregated score of 800.

Figure 12 displays the same analysis for a time slice selection of 18 minutes. Here, though, the most important factor is the Attack Maneuver Multiplier.

Once again we have uncovered something significant for validation of the model. Both sets of conclusions about what factor is most important to the outcome in this scenario cannot be correct.

We are therefore led to ask, "Which time slice is the correct one?" (or are both of them wrong). For a decision aid, should the analyst even have access to implementation details that can have such an impact? At the very least, the data farming effort made it easy to discover something about the model that deserves more attention.

SUMMARY OF FINDINGS

We found that Data Farming is an efficient way to find out what the model thinks is important. This finding speaks directly to the process of validation. For example, for the BBE tool, we discovered that time slice selection affected the relative performance of the COAs, changing the distribution of scores, changing which COA is "best," and changing which terrain modifier was most influential in the outcome. Such information is useful for software developers and users of the tool to consider.

CONCLUSIONS

The main interest Team 7 had coming into the workshop was to gain insight into how data farming might contribute to the validation process for a model. We immediately experienced first-hand the power of space-filling experimental designs when our first iteration highlighted the importance of the Road Speed multiplier in our test scenario.

A second, and more direct, contribution to our validation effort emerged when the highly exploratory nature of the data farming process allowed us to investigate model time-slice selection. Our discovery that the model changes character based on the selected time-slice is a significant finding that will be of immediate concern to the software developers.

Other data farming possibilities:

- Find the "right" time-slice value (or, why provide the selection?)
- Farm over COA parameters
- Farm over the value systems, such as the Commander's evaluation criteria weightings
- Tune performance of the decision aid by farming over Genetic Algorithm parameters to find the most efficient settings.
- Investigate sensitivities and impact of the global force modifiers provided in the decision aid (posture, C2, morale, etc.)

REFERENCES

[1] Schlabach, J. (2008). "BBE Cognitive Amplification for Battle Staff Planning." ERDC Topographic Engineering Center. Battlespace Terrain Reasoning & Awareness Battle Command (BTRA BC) Joint-Geospatial Enterprise Service (J-GES). Vol 3:2, pp 5-10.

[2] U.S. Army Field Manual 101-5. Staff Organization and Operations.

[3] U.S. Army Field Manual 34-130. Intelligence Preparation of the Battlefield.

[4] Petty, Mikel D. (2009). "Verification and Validation." In Principles of Modeling and Simulation: A Multidisciplinary Approach. Edited by John A. Sokolowski and Catherine M. Banks. John Wiley & Sons, Inc. Hoboken, NJ.

[5] Appleget, J., Duong, D., Blais, C., Perkins, T., and Brown, R. F. (2010). Irregular Warfare (IW) Model Validation and Best Practices Guide. DRAFT. Naval Postgraduate School. March 2010.

[6] Sanchez, S. M. 2005. "Work Smarter, not harder: Guidelines for designing simulation experiments," Proceedings of the 2005 Winter Simulation Conference, 69-82. Software and references available at http://harvest.nps.edu

[7] Old McData and other software and references are available at http://harvest.nps.edu

Plenary Sessions

Opening Session — 22-March-2010

- NPS OR Department Chair Welcome ....Chairman Robert Dell
- NPS Dean of Research Welcome ....Dean Karl van Bibber
- Keynote - Chronicle of Data Farming ....Klaus-Peter Schwierz
- HPC Forces Modeling and Simulation Overview ....Chris Bouwen
- DARPA Adaptive Execution Office Overview ....Ryan Paterson

Tuesday Plenaries — 23-March-2010

- Data Farming for New Members ....Gary Horne
- History of Data Farming................................Klaus-Peter Schwierz
- SEED Center Update and Design of Experiments......................Tom Lucas & Susan Sanchez
- Pythagoras...Donna Middleton
- Data Farming Tools ....Ted Meyer, Steve Upton, Mary McDonald

Wednesday Plenaries — 24-March-2010

- HPC PETT Program for Forces Modeling and Simulation ....Chris Bouwen
- A Brief Introduction to Modeling Emergent Behavior Systems ....Thomas Holland
- M&S of Human Factors: General Requirement for Data-Farming ....Harald Schaub
- Logistics Battle Command ....Jonathan Shockley
- O3158, an Experimental Agent Based Discrete Event Sim Framework..........................Steve Upton
- SASIO:Tactical Installation Protection.............Kenneth Byers
- Carrier Wing Tactics Incorporating Navy Unmanned Combat Air Systems ....Travis Gill
- JTB Data Farming for Test Planning ....Ed Lesnowicz

Team 8: Using Data Farming Techniques for Health Care Policy Analysis

TEAM 8 MEMBERS

Norman Reitter
Dan Widdis
Concurrent Technologies Corporation (CTC), US
Kah Wah Lai
Singapore

INTRODUCTION

Health care analysis has traditionally focused on understanding the impact of single intervention programs on single risk factors. Extensive research has been done on individual social risk factors that lead to disease. However, risk factors do not act independently. New research is required to understand the inter-relationships between environmental influences, social influences and human decisions across many risk factors. In addition, requirements are emerging to use a systems approach to analyze multi-factor intervention policies and the combined impact on overall population health and medical costs. This historical research approach and emerging needs have set up an environment that is ripe for using data farming techniques – large scale, efficient experimental design; creating data using modeling and simulation techniques; and a variety of statistical modeling methods to understand results.

In this paper, we discuss CTC's focus on using advanced analytical techniques for health care policy analysis. Specifically, we focus on progress made during the IDFW 20 in verifying an agent based simulation (ABS) model developed in a NetLogo® software program and refining a data farming approach for using the model for analysis.

Background

Over the past year, CTC, a non-profit scientific applied research and development corporation, began to develop an approach to use agent based simulation as part of a data farming approach to provide enhanced research for health care policy analysis. As part of its overall research effort, CTC has focused on a holistic solutions approach to providing systems analysis including:

- Research health care policy areas and integrate specific focus area data into formats usable for continued research
- Develop ontology models to represent the interrelationships between the human decision environment and influencers on human decisions (e.g. social networks, intervention programs, etc.)
- Develop an agent based simulation (ABS) model to represent interactions and assess future impacts of intervention policies on population disease rates
- Apply data farming techniques to the combined solution approach to analyze policies to support trade-off decisions.

Team Objectives

The IDFW 20 provided an opportunity for CTC analysts to verify the ABS model using the data farming techniques and leveraging SEED Center for Data Farming and IDFW 20 participant expertise to verify this approach and the ABS model. Pre-workshop objectives included:

- Develop an efficient experimental design to evaluate multiple health risk factors and understand their impacts on population health
- Use an agent based simulation model to harvest data for exploration and identification of potential intervention opportunities
- Evaluate the effects of single vs. multi-factor intervention policies on population health

Problem

CTC's research focused on answering the question "How do intervention policies impact population level characteristics?" The team focused on analyzing individual smoking characteristics, the impact of social networks to influence decisions to start or stop smoking, and the effectiveness of smoking intervention programs on reducing the overall population smoking rates.

Parameters and Measures of Effectiveness

For this part of our research, we use the percentage of smokers in the population as our main effect and six types of intervention programs:

- **ASPIRE** - Computerized smoking prevention curriculum: school-based self-study program
- **ESFA** - European Smoking prevention Framework Approach: integrated classroom with teacher, advertising, journalism
- **ASSIST** - A Stop Smoking in Schools Trial - school based, peer-led
- **PPBI** - Pediatric Practice-Based intervention - healthcare provider and peer-based
- **National Truth Campaign** - Advertising campaign and youth advocacy
- **SCYP** - Smoking Cessation for Youth Project

We vary intervention program coverage (or influence) over the population from 0% to 100%.

APPROACH

During IDFW 20, CTC's approach was to maximize use of data farming experts at the SEED Center and leverage expertise from other IDFW participants. To support this activity, prior to the workshop, the CTC team developed an ABS model using NetLogo® and collected data to support the model including population demographics and impacts of intervention programs on reducing smoking rates. We also calculated a variety of odds ratios for use in the ABS model and modeled the influence of an individual's social network on their chance of becoming a smoker or ceasing smoking.

The team's activities during the workshop, in figure 1, show how the ABS model evolved and a robust DOE developed.

Figure 1. Team 8 Activities during IDFW 20

As a result of our pre-work, we were able to run the ABS model using an initial experimental design on Monday. From this baseline, we were able to improve the model and the experimental design and make subsequent model runs. On Tuesday we enhanced the DOE to more fully explore the sample space. This activity allowed us to make another set of simulation model runs and analyze results starting on Wednesday. From this point, we were able to explore specific parts of the model in more detail to increase our understanding of the results.

ABS Model in NetLogo®

The CTC Team used the NetLogo®[1] software modeling language to develop an ABS model to support this project. Figure 2 shows a snapshot of the model version

In this model, each agent represented an individual. The individual agent had the following characteristics:

- Age
- Gender
- Race
- Relationships with other agents
- Smoker status (never, former, current)

Each agent maintained its smoker status at the age of 30 years for the remainder of its life. Life expectancy for each agent is based on actuary tables and current smoking status resulting in a chance that the agent dies each year based on current age and status.

The model uses a state-based probability of changes based on a set of odds ratios, developed from a significant amount of research from open source health research journals, to determine when an agent changes from one smoking state to another. The agent's social network, or set of peers, influences whether or not an agent changes smoking state.

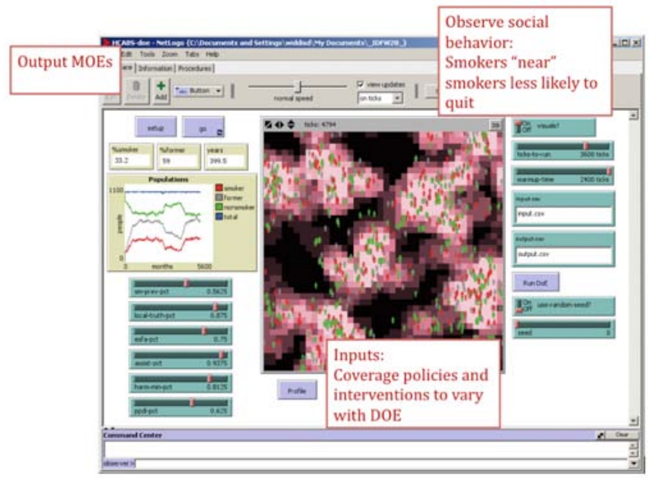


Figure 2. Picture Snapshot of ABS model in NetLogo® at beginning of workshop

The simulated environment lasts for approximately 300 years. We started the workshop considering 250 agents and expanded to 1000 agents by the end of the workshop.

Intervention programs were applied for a specified period of time over a portion of the population (from 0% to 100%), influencing the odds that an individual agent would change smoker states.

DOE Development

Our DOE evolved over the course of the week, resulting in a denser, robust examination of our desired sample space. Figure 3 shows how each iteration of our experimental design improved coverage and density of sample points within the design space.

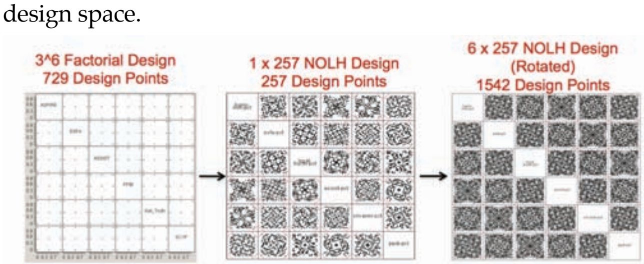


Figure 3. Evolution of Experimental Design

[1] http://ccl.northwestern.edu/NetLogo/

Our initial design used a full factorial considering each of the six interventions at three levels (0, .5, 1) resulting in 729 design points, or simulation model runs. We improved our design by using a Nearly Orthogonal Latin Hypercube (NOLH) model based on the SEED Center NOLH spreadsheet.[2] This provided a more robust sampling of our design space, with a reduction of 257 design points, however continued to expose some gaps based on the resulting combinations of intervention coverage inputs. We were also concerned that results showed possible dependencies between interventions-based sampling patterns. After conferring with SEED Center staff on how to more completely fill our design space, we used a rotated NOLH design, resulting in 1542 design points (simulation runs) and much richer sampling space represented by the right hand side design in figure 3. This final DOE allowed us a robust and efficient sampling plan to examine all combinations of intervention programs at many different levels (ranging from 0…1).

RESULTS

Throughout the workshop we compared the percentage of smokers within a population prior to applying the interventions and then once the population smoking percentage reached a steady state after the interventions were applied over a period of time.

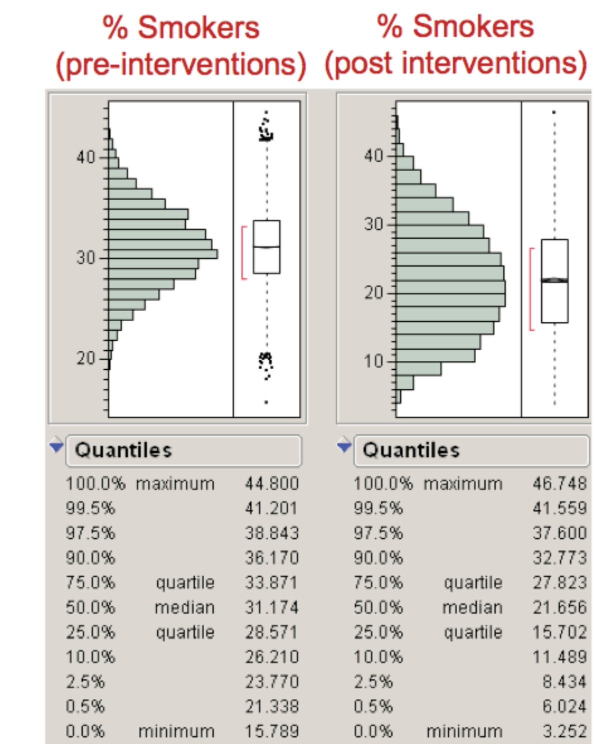


Figure 4. Example Model Output: % Smokers Pre and Post Interventions

Figure 4 shows an example of the distribution of percentage of smokers before and after the interventions were applied. In this example, the smoking population shifted ten percentage points from 31% to 21% of the total population.

With the effectiveness of each intervention being fundamental to our research, we next evaluated how each intervention acted independently on the reduction in the percentage of smokers in a population. In figure 5 we show an example of how one of the interventions, in this case the ASSIST intervention program, dominates the other interventions shown by a significantly steeper positively increasing slope when varied over increasing levels of coverage across the population (from 0% to 100%) with the most impressive influence when applied to over 50% of the population.

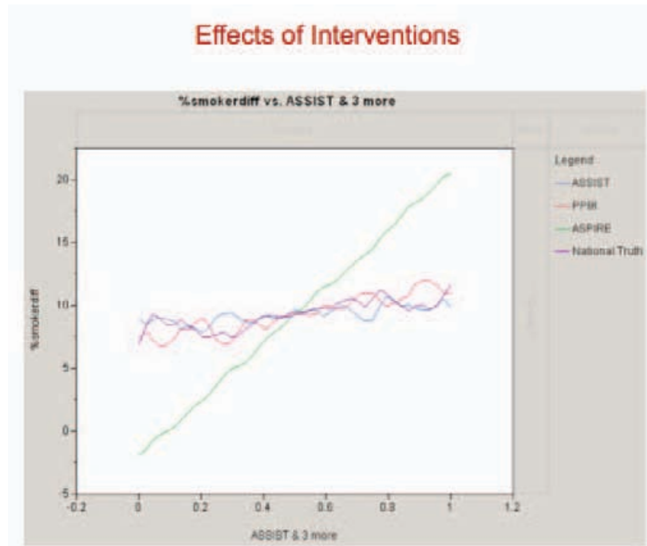


Figure 5. Example Model Output: Relative Effects of Intervention Programs

The other interventions, in this example PPBI, ASPIRE, and National Truth Campaign, are more effective then the ASSIST intervention program at a level up to 50% population coverage, however, the ASSIST intervention then dominates the other interventions. Using this example, if a policy maker only had enough funds to invest in a program that influenced up to 50% of the population, we would recommend that they chose any of the programs except the ASSIST program. However, if they had enough funding and a desire to implement an intervention program over more than 50% of the population, then the ASSIST program is a much more effective choice for reducing population smoking rates.

As part of our model assessment, we decided to compare the performance of our agent based simulation model in predicting smoking rate reduction vs. how the odds ratios predicted the same outcomes. Figure 6 shows a comparison of how the simulation output (plot points on the graph) compare to our estimate of population smoking reduction based on the odds ratios (linear line plot). Most of the odds ratios performed as expected with the exception of the Truth Campaign intervention, which led us to hypothesize, that the original research to support the effectiveness of this intervention may not have considered the effects of social

[2] See software downloads at http://harvest.nps.edu/

influences on the outcome. This area is one that we will explore with further research.

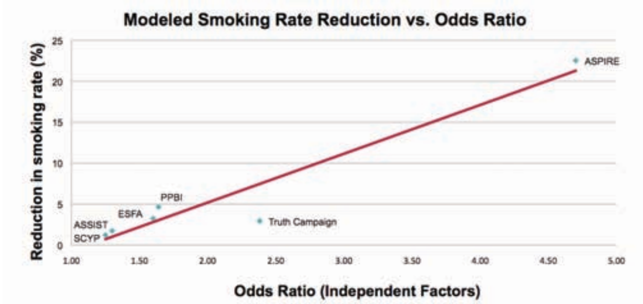


Figure 6. Example Model Output: % Smokers Pre and Post Interventions

Next, we explored how well the agent based simulation model showed how the population behaved over time based on the human and intervention program characteristics.

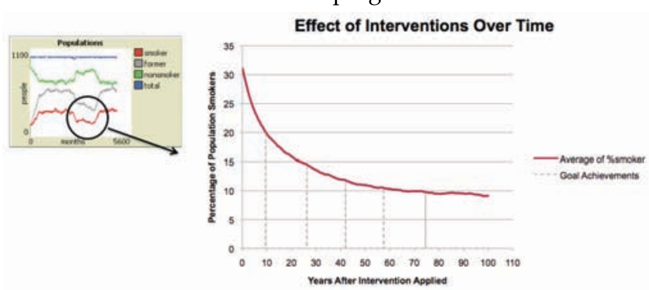


Figure 7. Example Model Output: Time for Interventions to take effect

Figure 7 shows how the model behaved for an example combination of interventions. In this case, the model projected approximately 60 years to reduce the population smokers by 20%

CONCLUSION

Leading into IDFW 20, Team 8's objectives centered on developing a robust experimental design and conducting verification and limited validation of the agent based simulation model that we developed using NetLogo©. During the workshop, we accomplished all of these objectives and realized tremendous improvements through the help of solid preparation, other IDFW 20 participants, and the expertise of the SEED Center for Data Farming professionals.

Special thanks goes to Santiago Balestrini, another workshop participant, for selflessly providing his time and NetLogo© knowledge to help improve our model run time. Through improvements to our experimental design and agent based simulation model, we are now able to explore a more robust sample space with ½ of the original model run time resulting in a more robust analysis capability.

Our initial technical observations as a result of this workshop include:

• Validation that the influences of a social network are important to consider when evaluating the effectiveness of intervention programs on reducing population smoking rates

• An estimate for the length of time each intervention program or combination of programs need to be funded to ensure effective reduction in population smoking rates

• An understanding of which type of intervention(s) to invest in based on the size of the population that can be reached based on limited funding.

Following this workshop, CTC will build on the insights gained during IDFW 20 by continued enhancement of the agent based simulation model. This work supports a larger research effort to support policy decisions that effect funding for different types of intervention programs based on expected effectiveness in reducing smoking rates with extension to disease prevention.

Team 9: Evaluation of a USMC Maritime Prepositioning Force Arrival and Assembly Model

TEAM 9 MEMBERS

Jonathan Derosier, Captain
U.S. Marine Corps

Hans Neumueller, Lieutenant Colonel
German Army

Helcio Vieira, Major
Brazilian Air Force

Jamie Breakfield
Northrop Grumman Shipbuilding, US

Edward Lesnowicz, Colonel, USMC (ret)
Dr. Keebom Kang
Naval Postgraduate School, US

INTRODUCTION

The U.S. Marine Corps' Maritime Prepositioning Force (MPF) enables the rapid deployment of Marine forces to permissive areas of operations. The MPF consists of more than a dozen ships divided between three squadrons. Each squadron supports a notional Marine Expeditionary Brigade (MEB) and is based in one of three locations: the Pacific Ocean, the Indian Ocean, or the Mediterranean.

MPF Operation

During an MPF operation, a Maritime Prepositioning Ship Squadron (MPSRON) or some portion or combination thereof, is deployed to a permissive area of operations where its equipment and supplies are offloaded. A fly-in echelon (FIE) compromising the bulk of personnel and additional equipment is flown into a nearby airport. The equipment and personnel are then integrated to form a functioning Marine Air Ground Task Force (MAGTF). This process is called Arrival and Assembly.

Motivation

Operations Enduring Freedom and Iraqi Freedom (OEF/OIF) have caused rapid modernization of the USMC's equipment systems since 2003. This equipment is now being incorporated into the MPF program with potential impacts on Arrival and Assembly. An example is the armoring of the Medium Tactical Vehicle Replacement (MTVR), which is 'reduced' for embarkation on ship. During Arrival and Assembly, the MTVR armor needs to be reconfigured; a three-hour process requiring two mechanics and a piece of Material Handling Equipment (MHE) with its operator.

The tradeoff between resources (mechanics, container handlers, etc.) and the force generation timeline during MPF Arrival and Assembly is of particular interest.

Analytical Framework and Goals

An analytical framework is illustrated in Figure 1. The goal of this work at IDFW 20 is to use data farming techniques to analyze an MPF Arrival and Assembly model to inform data collection efforts for future MPF operations and/or exercises.

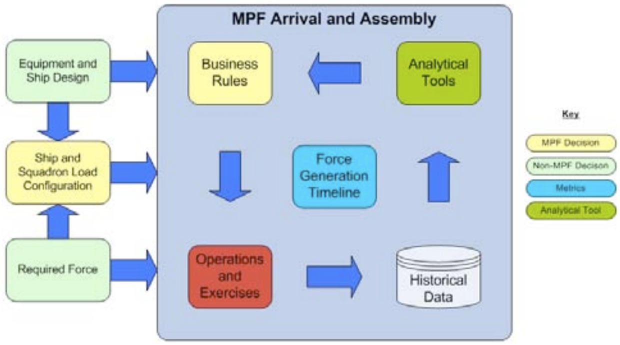


Figure 1. MPF Arrival and Assembly Analytical Framework

ARRIVAL AND ASSEMBLY MODEL

The MPF Arrival and Assembly Model is a discrete event simulation implemented in ExtendSim7. The model has two main processes: the offload of equipment from a ship to a pier and the throughput of equipment from the pier to its using unit located some distance from the pier.

Offload

The offload process models the interaction between ships and docks, where a dock is required to conduct an offload. Multiple docks allow for the simultaneous offload of ships. There are two methods for offloading equipment from a ship:

1. Roll On Roll Off (RORO) is used for vehicles that can be driven off the ship via its stern ramp. RORO requires both a ramp (ship asset) and offload drivers.
2. Lift On Lift Off (LOLO) is used for offloading containers (and possible vehicles) by lifting them with either a ship crane (ship asset) or a gantry crane (dock asset).

All the equipment is offloaded in a random order from the ship with all vehicular equipment using RORO and all containerized equipment using LOLO.

Throughput

The throughput process models the physical movement of equipment from the pier to the using unit and any maintenance or setup actions that must be completed to make equipment operational. The equipment is classified by type with each type requiring various assets during throughput as identified in Table 1.

Throughput Assets	Equipment Type				
	AMMO	ISO	RS	MTVR	HET
Throughput Driver			X	X	
Mechanic			X	X	X
RTCH, Pier	X	X			
RTCH, CSA	X	X			
Armor Teams				X	
Truck		X			
Truck Convoy	X	X			
Driver Convoy			X	X	
HET Convoy					X
Security Convoy	X	X	X	X	X

Table 1: Throughput Asset and Equipment Dependencies

The equipment must undergo various sub-processes dependent upon equipment type and additional factors such as a piece of equipment being 'frustrated' (dead lined and requiring maintenance) at the time of offload. The dependencies between throughput assets and sub-processes are identified in Table 2.

Model Parameters

Each of the offload and throughput resources is a parameter that can be controlled in the model. Additional parameters are:

1. Ship Crane Delay - time required to offload a piece of equipment using a ship crane.
2. Gantry Crane Delay - time required to offload a piece of equipment using a gantry crane.
3. Ramp Delay - time required to offload a piece of equipment using a ramp.
4. Return Offload Driver Delay - time required for an offload driver to return to the ship and be available to offload another vehicle.
5. Rough Terrain Container Handler (RTCH) Delay - time required to load/unload a container on a truck.
6. Truck Speed - speed at which a truck for moving ISO containers moves within the port.
7. Pier to Container Storage Area (CSA) Distance
8. Frustrated Delay - time required for a mechanic to repair a frustrated piece of equipment.
9. Frustrated Rate - probability that equipment is frustrated at offload.
10. SL3 Delay - time required to set up SL3 equipment on vehicles.
11. Mechanic Priority - the relative priority of SL3 vs. frustrated equipment for mechanics.
12. MTVR Armoring Resources - the number of resources dedicated to armoring MTVRs.
13. Rolling Stock (RS) to Movement Control Center (MCC) Delay - time required to move RS vehicles from the pier to the MCC staging area where they are formed into convoy sticks by destination.
14. Port to Destination Distance - distance from the port to the final destination. Each destination is an independent variable.
15. Convoy Delay at Destination
16. Convoy Priority - priority for assigning security assets to convoys.

Scenario

In this scenario, we model a single MPSRON offload. The MPSRON has 4,298 Principle Equipment Items (PEIs) spread across four ships with the following breakdown by equipment type:

- ISO (General Cargo Containers) 42%
- RS (Rolling Stock Vehciles) 32%
- AMMO (Ammo Containers) 14%
- MTVR (Sub-set of Rolling Stock) 7%
- HET (Tracked Vehicles) 5%

Metrics

Figure 2 is a screen shot of the model outputs. The blue, green and red lines represent the counts of equipment over time at the pier, at the final destination, and in the throughput process respectively. We use days to complete offload, the days to complete throughput and the mean cycle time of equipment (time complete - time offloaded) as our primary metrics.

Throughput Assets	Throughput Sub-Processes					
	Frust-rated	SL3 Setup	Move ISO Pier to CSA	Move RS Pier to MCC	Armor MTVRs	Convoy To Using Unit
Throughput Driver				X		
Mechanic	X	X				
RTCH Pier			X			X (Ammo)
RTCH CSA			X			X (ISO)
Armor Teams					X	
Truck			X			
Truck Convoy						X
Driver Convoy						X
HET Convoy						X
Security Convoy						X

Table 2: Throughput Assets and Sub-Process Dependencies

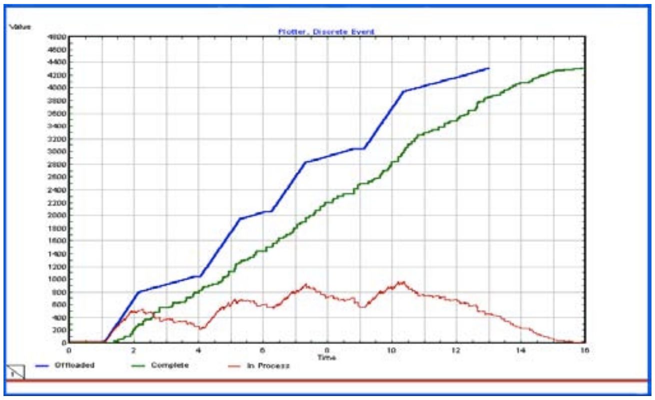


Figure 2. Simulation Output. The blue line identifies the count of equipment as it is offloaded at the pier. The green line is the count of equipment as it arrives at the final destination. The red line is the count of equipment in the throughput process. The primary metrics in the simulation are the day offload is completed, the day throughput is completed, the mean flow time (time complete - time offloaded) of equipment in the throughput process.

Factor	| Effect |
gcrane	192.60638
dock	174.58188
dock:gcrane	73.28010
scranedelay	39.67487
gcrane:scranedelay	35.61884
gcranedelay	21.38376
returndelay:offloaddriver	15.53665
dock:scranedelay	14.69414
dock:gcranedelay	13.08632

Table 3. Top nine factors including two-way interactions that effect the time to offload

DESIGN OF EXPERIMENTS

We use a 28-factor Nearly Orthogonal Latin Hypercube (NOLH) design of experiments with 200 design points. Each design point was run 30 times for 7,200 total runs.

RESULTS

Offload

The use of gantry cranes and the number of docks has the highest impact on the offload completion time as shown if Table 3. Given that 56% of all equipment items in the model are containers offloaded by cranes and a gantry crane is much faster then a ship's crane this result is not a surprise.

Throughput - All Equipment

Of the 7,200 runs, the proportion of runs where a particular equipment type was the last to arrive at its final destination has the following break down:

- ISO 87%
- HET 3%
- RS 1%
- MTVR 7%
- MTVR/RS 2%
- AMMO <1%

The MTVR/RS are cases where both the MTVR and RS were completed at the same time. This situation occurs because MTVRs are a subset of RS and both may travel in the same convoy. All pieces of equipment in the same convoy arrive at the destination at the same time.

The nine most significant factors that affect the final destination arrival time are listed in Table 4. Of these, the top eight factors are directly related to the throughput of containers. Considering that ISO containers finished last during 87% of the simulation runs this result is not surprising.

Factor	| Effect |
rtchdelay	407.9187
rtchcsa	401.8970
truckconvoy	212.9060
securityconvoy	202.0804
rtchpier	126.8719
aaoedistance	119.8650
gcranedelay	118.0914
dock	102.9512
sl3delay	100.9018

Table 4. Top nine factors including two-way interactions that effect the latest time for all equipment to arrive at its final destination(s).

The AMMO containers almost never finish last because they are formed into convoys directly on the pier and they have the highest priority when assigned convoy security.

Throughput - By Equipment Type

The overall time to complete the throughput does not paint a complete picture because it is highly influenced by the ISO containers. It is reasonable that equipment types with proportionally more equipment will take longer to throughput than those with proportionally lower equipment. In addition, individual equipment types use different sub-processes and resources during the throughput process.

Table 5 shows the top five factors that effect the throughput completion time of each equipment type. This table illustrates the ranking of factors across the equipment types.

First, it is clear that RTCH plays a significant role in the throughput of containers. Both the delay and the number of RTCHs are significant for ISO containers and the number of RTCHs at the pier is significant for the AMMO containers (Note 1).

The throughput of AMMO containers is affected more by the container offload rate (Note 2). This result occurs because the AMMO containers have the simplest throughput process as they are convoyed directly from the pier.

However, the AMMO containers are still affected by the number of security assets, as are all other equipment types (Note 3). Security plays a particularly strong role across the three vehicle equipment types.

The convoy transportation assets have high-ranking effects for each equipment type except AMMO containers (Note 4).

AMMO	ISO	RS	MTVR	HET
gcrane[2]	rtchdelay[1]	securityconvoy[3]	securityconvoy[3]	securityconvoy[3]
dock[2]	rtchcsa[1]	sl3delay[5]	gcrane[6]	hetconvoy[4]
securityconvoy[3]	truckconvoy[4]	rsconvoy[4]	sl3delay[5]	sl3delay[5]
rtchpier[1]	securityconvoy[3]	gcrane[6]	rsconvoy[4]	gcrane[6]
dock:gcrane[2]	rtchpier[1]	mechanics	dock:securityconvoy	mechanics

Table 5: Top five factors that impact the throughput completion time by equipment type. Notes for the factors are indicated by superscripts and explained below.

Note 1: The number of RTCHs and the RTCH Delays are highly ranked for the throughput of containers, particularly ISO containers. AMMO containers are less affected because they are convoyed directly from the pier.

Note 2: The dock and use of gantry cranes are highly ranked for the throughput of AMMO containers. AMMO containers have simplest throughput process as they are convoyed directly from the pier. Therefore, the offload rate has a larger effect on their throughput process than for that of other equipment types.

Note 3: The number of security assets is highly ranked across all equipment types and particularly for vehicles.

Note 4: The number of convoy assets is highly ranked for each equipment type except ammo containers.

Note 5: The SL3 setup delay is highly ranked for all of the vehicles.

Note 6: The use of gantry cranes is highly ranked for the vehicles, which do not explicitly use the gantry cranes. This situation may be because of the dependency created by the sharing of the security assets.

The SL3 setup delay is ranked highly across the vehicle equipment types (Note 5).

Surprisingly, MTVR armoring resources ranked 94th among the MTVR throughput completion time factors. This indicates that the armoring process is relatively unimportant to the final MTVR throughput completion time. However, the correlation between the proportion of runs where MTVR throughput completion is greater than or equal to RS throughput completion is -0.925 indicating that increasing the armoring throughput (more armoring resources) has an important effect on when the using units will receive their MTVRs relative to other RS vehicles.

MTVR Armoring Throughput	Completion of MTVR >= RS		Percent TRUE
	FALSE	TRUE	
1.397	3	447	99.3%
2.2195	51	849	94.3%
2.794	68	832	92.4%
3.6165	132	768	85.3%
4.191	139	761	84.6%
4.439	207	693	77.0%
5.0135	203	697	77.4%
5.836	195	705	78.3%
6.6585	101	349	77.6%

Table 6. The correlation between the proportion of runs where MTVRs throughput completion time is greater then the RS throughput completion time and the MTVR armoring throughput rate is -0.975.

Finally, the use of gantry cranes is ranked high for the RS and HET required vehicles (Note 6). This result implies that the offload rate of containers is somehow affecting the completion time of these two equipment types. The only cross dependency between the vehicles and the containers is via the convoy security assets.

Throughput - Flow Time

The factors that are significant to the mean equipment flow time (the time an item arrives at the final destination - the time it was offloaded) are listed in Table 7.

Container handling (RTCH Delay, number of RTCHs) has the largest effects. Considering that ISO containers comprise 42% off the equipment and every ISO container is touched three times by a RTCH during the throughput process it is not surprising that these factors have large effects.

Factor	| Effect |
rtchdelay	108.19988
rtchcsa	87.33917
securityconvoy	86.74715
sl3delay	56.81665
rtchpier	52.65566
mechanics:tpdrivers	49.61680
rtchdelay:rtchcsa	39.91947
mechanics:securityconvoy	37.93413
rtchpier:rtchcsa	37.71522

Table 7. Top nine factors including two-way interactions that effect equipment flow time

SUMMARY AND WAY AHEAD

This evaluation of the MPF Arrival and Assembly Model has identified key parameters and processes in the model that have high effects on the model's measures of effectiveness (throughput completion time, equipment flow time, and offload completion time). The most important factors and processes in the model are:

- The handling of containers including RTCH delays and the number of RTCHs in use.
- The number of convoy transportation and security assets.
- The use of gantry cranes or not.
- The SL3 setup delay.

These factors and processes should be the focus of future MPF exercise data collection efforts.

Additionally, future work on the model should focus on validating that the real world processes that are most significant to the model are adequately and accurately represented. For example, the convoying of equipment is currently grouped by both equipment type and destination. It may be more appropriate to have equipment of varying type but the same destination travel in the same convoy.

Team 10: Robustness Analysis with ITSim: Patrol Planning and Plan Execution Simulation

TEAM 10 MEMBERS

M Thomas Orichel
IT-AmtBw, Germany

Christoph Mies
Benjamin Wolters
Sascha Hörcher
Fraunhofer IAIS, Germany

LtCol Dietmar Kunde, PhD
HFüKdo, Germany

Kurt Grau
IABG, Germany

Abstract

ITSim is a general purpose simulation system for decision-support. It focuses on the simulation of coherent processes and provides additional methods for examining optimization tasks within the broader range of tasks of the German Armed Forces, the Bundeswehr. Modern warfare scenarios are dominated by asymmetric threats with complex non-linear interdependencies and interrelations that traditional techniques of analysis are insufficient to capture.

For example, it is hard to determine the cost and benefit of force deployment at several bases in the area of operation (AOO). On the one hand, the deployment at several bases has a positive effect because the forces are spatially closer to the points where mission objectives have to be accomplished. On the other hand, longer supply chains have to be guarded. IT-AmtBw and Fraunhofer IAIS are currently developing an extension to ITSim that provides decision support on this optimization problem. Several factors are involved in such an investigation. One core factor is the generation of a patrol plan, which is a schedule for all designated forces for a certain time horizon. It maximizes the presence at certain points of interest (POI). A POI is an element of a mission that requires special actions, like reconnaissance, presence, show of forces or CIMIC activities. It is formalized as a location, desired visit frequency, a certain duration and a weight. Thus, a POI located at location is to be visited regularly with a time interval of frequency. Each visit lasts duration time units. The relative importance of one POI with respect to the other ones can be modeled by assigning weight as a multiplier.

The overall questions to be answered with an investigation like this one are the following:

- Is it better to distribute forces to several bases, or instead to concentrate them in one single base ?
- If the forces are to be distributed, which distribution is optimal ?

Note that we cannot answer these questions at this early stage of development. Additionally, many factors are important for a well-founded estimation. At this workshop, we wanted to answer the following questions:

- Can ITSim find optimal patrol plans with respect to the patrol presence at given POIs?
- Is the given (technical) concept of patrol presence also suitable for military definitions of patrol presence?
- What is the influence of weighted, i.e. prioritized, POIs?
- How robust are the generated patrol plans against execution flaws?

The following two sections introduce the scenario at hand as well as the performed analyses and their results. The investigation is divided into two phases, optimization and simulation. Both parts are discussed in the sections. The final section gives a conclusion and suggestions for future work.

Figure 1: Scenario Overview

Scenario

Figure 1 depicts the investigated scenario, which is located in Germany. The area has a size of roughly 35 km times 29 km. Three bases, marked with a red circle, are located in the scenario. The other icons represent POIs, which have to be included in the patrol plan with a certain time of patrol presence.

Figure 2 depicts another view at the scenario. The roads (dark lines) are imported from a data source from the German Armed Forces. The POIs are categorized into several classes: red, yellow and green. To each class different parameter values are assigned. The analysis of the scenario is divided into two phases, an optimization and a simulation phase, both discussed in the following.

Figure 2: Roads and classification of POIs

Optimization Phase

During this phase, an optimal schedule (patrol plan) for all patrols is generated. This optimization is a hard problem. The patrols are constrained by their fuel capacities as well as by a maximal operational duration per day, which must not be exceeded. In our experiments, the patrol plan has a time horizon of 20 days, i.e. each patrol is assigned tasks for 20 days. Every patrol must return to its home-base every evening in order to rest and re-supply. All patrols have the same average speed of 40 km/h. In Table 1, the parameters of the different classes of the POIs are listed.

Class	Frequency	Duration
Red	12 h	1 h
Yellow	24 h	2 h
Green	36 h	3 h

Table 1: Parameters of the different POI classes

After the patrol plan has been generated, it can be integrated into the scenario and its execution can be simulated.

Simulation Phase

During the simulation phase, the robustness of the generated plan is analyzed. This analysis is important since there are always discrepancies between operation planning and operation execution. Thus, the best plan during optimization might not be the best plan during execution.

In order to disturb the plan execution, we have defined some unexpected events that occur stochastically: Blocked roads and stochastic patrol durations. In order to enforce re-routing of patrols, some roads are blocked stochastically. Thus, the patrols cannot take the shortest route and the required time to reach a POI or a base increases. The second event reflects the effect that mission execution at a certain POI is not always straight-forward as expected. Delays as well as unexpected fast accomplishment can occur. Thus, the POI duration is not deterministic but stochastic.

The Design of Experiment (DoE) as well as the results are introduced in the next section.

Results and Analysis

In this section, we introduce the conducted experiments and present the analysis which is used to answer our questions stated in the introduction. The results are presented according to the two phases mentioned above.

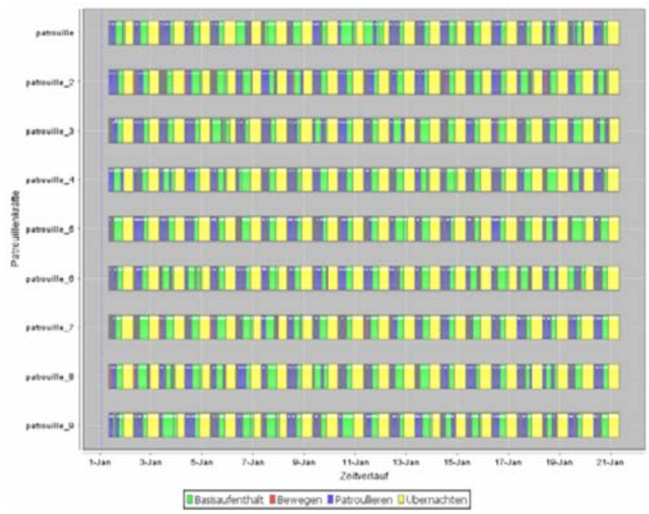


Figure 3: Patrol-centric view on a patrol plan

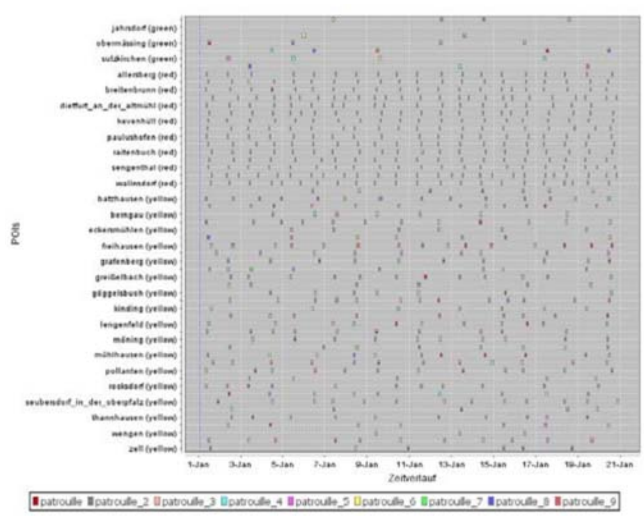


Figure 4: POI-centric view on a patrol plan

The first two questions were discussed in detail at the workshop. We implemented many views on the resulting plans, e.g. Figure 3 and Figure 4, in order to analyse the quality of each generated plan. The former figure shows the actions of the patrols in time (green: at base, red: move, blue: patrol, yellow: sleep). The latter shows the visits of the different patrols at each POI over time.

Our discussions revealed that the technical concept of patrol presence is not sufficient for a military decision maker. The reason for that is that there are many, often conflicting, goals to be pursued. We will extend ITSim in future upgrades to overcome this shortcoming by introducing more optimization criteria and performing a multi-dimensional optimization. One additional but not sufficient criterion is introduced in the following and will be compared with the current technical criterion.

Optimization Results

The results of the optimization are discussed below. We first generated a patrol plan according to the scenario depicted in the figures above. Seven patrols are distributed over the three bases. The influence of different weights for the POI classes should be determined. Therefore, we calculated the number of POIs which have been satisfied, i.e. where the desired patrol frequency is never violated. When the POI is visited at a certain point t in time, the next visit should happen exactly at point t', which is t plus the duration of the visit and the desired frequency time. When the next visit happens at t' plus minus a certain tolerance value, the next visit is in time. The tolerance value is a percentage of the frequency of that particular POI. A POI is satisfied if all visits are in time.

Three different weight combinations for the three POI classes (red, yellow and green) have been selected. We call them single, double and triple weighting. In the first case, all classes have the same weight, in the second case red is twice as important as yellow which is again twice as important as green. In the last case, the weights differ by the factor three.

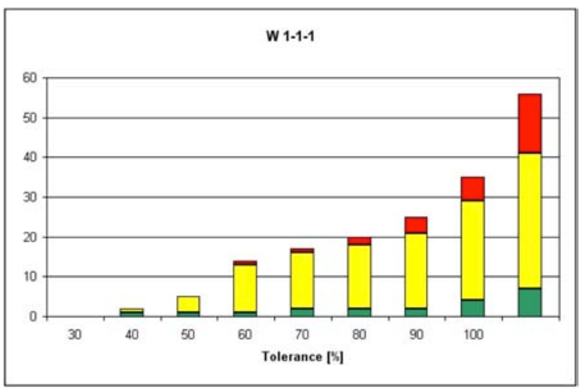


Figure 5: Satisfied POIs with 'single weighting'

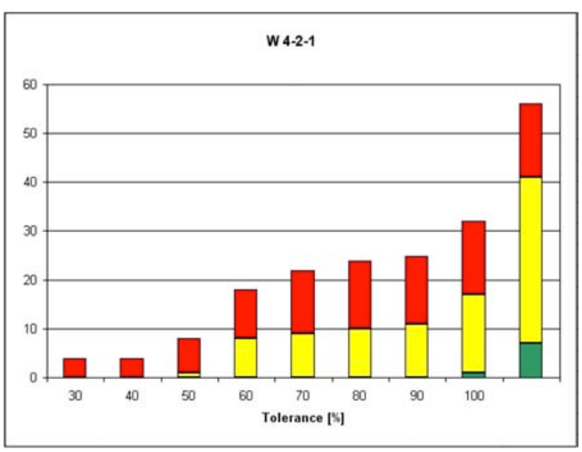


Figure 6: Satisfied POIs with 'double weighting'

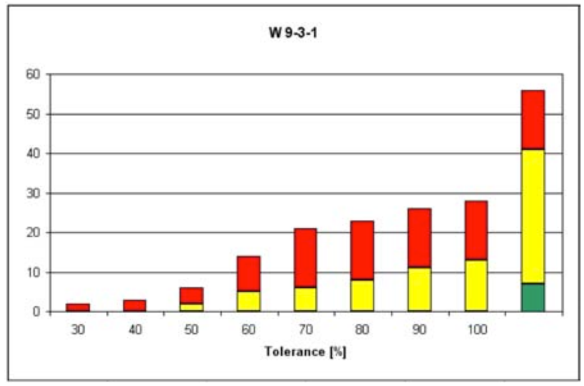


Figure 7: Satisfied POIs with 'triple weighting'

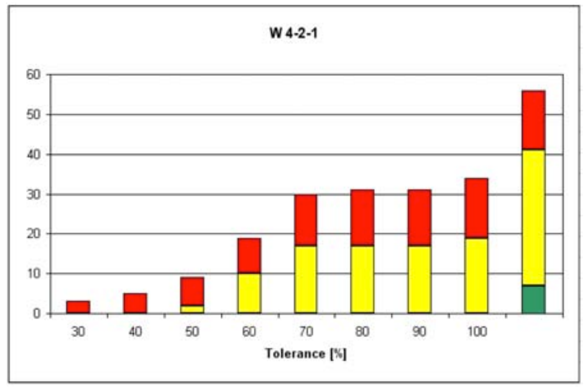


Figure 8: Satisfied POIs with 'double weighting' with one base

Figure 5 to Figure 7 depict the results of the experiments. The last bar always represents the overall number of red, yellow and green POIs, respectively. The higher the difference in weight, the more red POIs are satisfied. At the same time, the number of overall satisfied POIs decreases since the patrols are concentrated on the important POIs and do not take much care about the unimportant ones. Thus, the user has to carefully select its prioritization.

The aim of the next experiment is to compare the deployment to three bases with a deployment to one base, namely 'fob_großalfalterbach', the top-right base depicted in Figure 1. Since only one base has to be defended, two more patrols, namely nine, can be employed for patrolling. We performed one experiment with 'double weighting'. The results for one base are depicted in Figure 8 and can be compared with Figure 6, where three bases were used.

Tolerance	Green		Yellow		Red	
	1	3	1	3	1	3
30%	0	0	0	0	3	4
40%	0	0	0	0	5	4
50%	0	0	2	1	7	7
60%	0	0	10	8	9	10
70%	0	0	17	9	13	13
80%	0	0	17	10	14	14
90%	0	0	17	11	14	14
100%	0	1	19	16	15	15

Table 2: Number of satisfied POIs with 'double weighting' for one and for three bases

Table 2 shows the number of satisfied POIs clustered into their class with respect to a given tolerance. The numbers are calculated for the deployment to one and three bases. We can see that the deployment is superior if only one base is used. The green class with 100% tolerance is the only outlier. This is also confirmed by our technical optimization criterion, the patrol presence. The patrol plan with seven patrols and three bases gained a value of 1613, wheras the patrol plan with nine patrols in one base realized a value of 1654. Nevertheless, it seems reasonable to use a multi-dimensional optimization method in the future.

Simulation Results

For the simulation of the patrol plan execution, we used the best plan of the double weighting setting with one base. As mentioned above, the unexpected events were road blocks and stochastic patrol durations at the POIs. Additionally, we varied the speed of the patrols.

Factor	Min	Max	Unit
Patrol Speed	20	60	km/h
Blocked Roads	0	500	
Block Duration	1	48	Hour
Block Start	0	19	Day
POI Duration	20	180	Minute

Table 3: NOLH design

Table 3 contains the *Nearly Orthogonal Latin Hypercube (NOLH)* [1] design of our experiment. Every parameter configuration has been run with *17* different seeds. The simulation part revealed that our plans are very robust since no varied parameter has a statistically significant impact on the regarded *Measure of Effectiveness (MoE)*, which is the number of satisfied red, yellow and green POIs. The main reason therefore is probably that all delays are compensated by the nightly rest. Additionally, the number of closed roads was probably too small. We have to invest more time in order to evaluate the robustness of the generated plans in more detail.

CONCLUSION

Intelligent force deployment is a difficult optimization problem. Many, sometimes conflicting, critera influence the final decision. The patrol plan generation module of ITSim, which is currently still under development, might support a human decision maker, i.e. the commanding officer.

In our goal to analyze the robustness of the patrol plans we foccussed at first on the plan generation itself and discussed appropriate quality measures for plans. A broadly accepted notion of an optimal patrol plan is very hard to develope since it is always subject to the current situation and main intent of the decision maker. One way out of this dilemma is to integrate many different possible critera and optimize them simultaneously in a multi dimensional optimization (e.g. [2,3]). Afterwards, the decision maker can select among the solutions and adjust the tradeoffs manually.

Another very important idea for future work is the investigation of the tuning of technical parameters of the optimization in ITSim. Because a genetic algorithm is employed, many parameters are used to adjust the search heuristic, i.e. the genetic operators. Perhaps a more optimal parameter configuration can be found automatically.

REFERENCES

[1] T. M. Cioppa, “Efficient nearly orthogonal and space-filling experimental designs for high-dimensional complex models,” PhD Thesis, Naval Postgraduate School, Monterey, CA, 2002

[2] K. D. Deb, A. Pratap, S. Agarwal and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm : NSGA-II," in IEEE Transactions on Evolutionary Computation, 2/2002, pp.182-197

[3] E. Zitzler, K. Giannakoglou, D. Tsahalis, J. Periaux, K. Papailiou and T. Fogarty, "SPEA2: Improving the Strength Pareto Evolutionary Algorithm for Multiobjective Optimization," TechReport, 2002

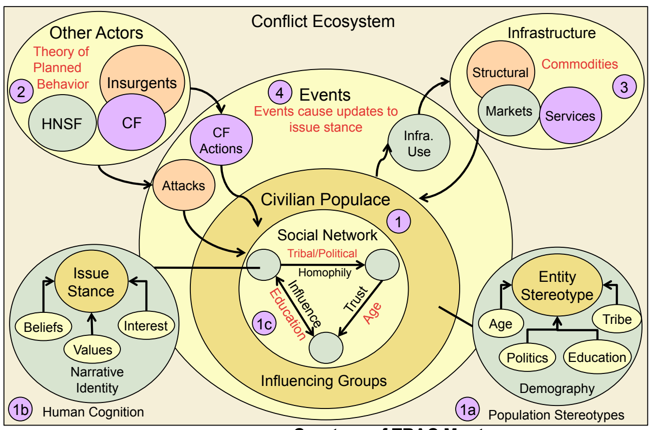
Poster Session

Team 11: Representing Violent Extremist Networks within Social Simulations for Attack the Network Course of Action Analysis

TEAM 11 MEMBERS

MAJ Jon Alt
MOVES, Monterey, USA

LtCol (HEA) Sotiris Papadopoulos
MOVES, Greece



Courtesy of TRAC Monterey

Figure 1. Cultural Geography Model

Overview

The Cultural Geography (CG) model, shown in Figure 1, is a government owned, open source prototype agent-based model of civilian populations currently implemented in Java and using Simkit as the simulation engine.

The model aims, through the implementation of social and behavioral science, to track individual, group-level and population-wide changes on positions related to various issues.

At its current stage the model examines the issues of security, elections and infrastructure.

Goals

We had the following goals for IDFW 20:

- Create an agent prototype that decides on its actions using utility theory.
- Create code to support the utility agent's decision process.
- Test the utility agent's functionality within the CG model.
- Design an experiment using Data Farming techniques for evaluating the utility agent's performance

Analysis

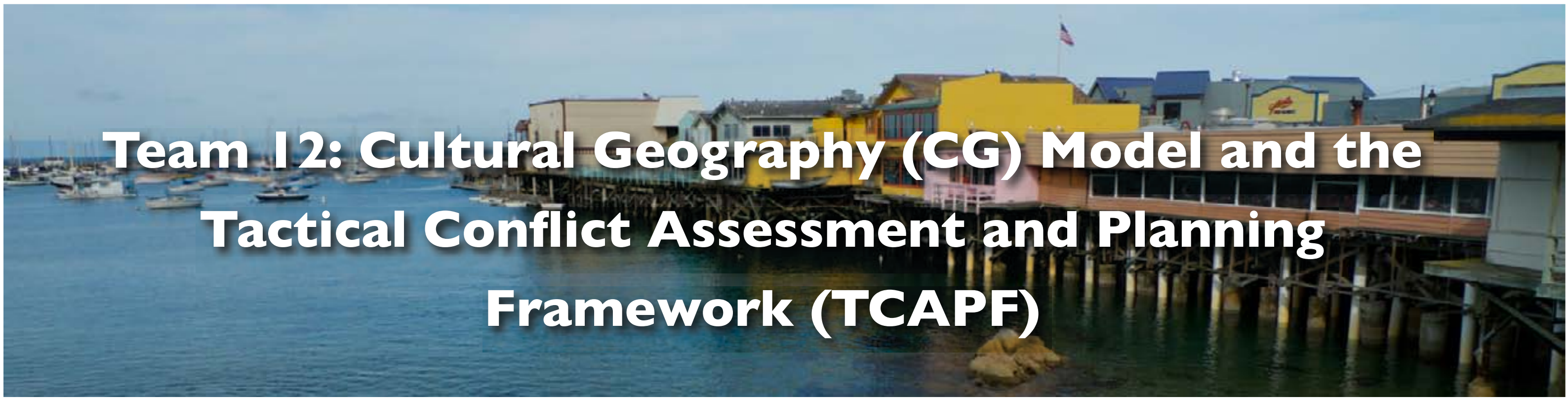
Our methodology include the development of an Agent Template for implementation, improvement and finalization of the template, development of an experimental design, and comparative analysis with different utility functions and roles.

The principle of maximum expected utility (MEU) says that a rational agent should choose an action that maximizes the agent's expected utility. For the purposes of this project, we consider as utility the change in the population's stance on the issue of Security.

To determine the utility of an action we tracked the execution of each action, track the utility accumulated following each rule firing, discounted the utility to determine the present value of the utility at the time of execution, and determined the mean utility received for each rule fired at the time of firing. We then selected action based on the activation level and the Boltzmann distribution. Our initial violent extremest network consisted of 30 insurgents across 10 zones.

Future Work

Our plans for future work include incorporating additional attributes into the utility function, developing additional roles within the insurgent network, and exploring the use of different utility functions for different roles within the network.



Team 12: Cultural Geography (CG) Model and the Tactical Conflict Assessment and Planning Framework (TCAPF)

TEAM 12 AUTHORS

Perkins, Timothy K.
U.S. Army Engineer Research & Development Center

Pearman, Gerald M.
Augustine Consulting, Inc., US Army Retired

Bachl, MAJ Christopher A.
Naval Postgraduate School, US Army

Vargas, MAJ Joe A.
TRADOC Analysis Center – Monterey, US Army

INTRODUCTION

A goal of stability operations is to influence civilian attitudes in favor of the host nation (HN) government and the stabilization forces. To help understand the dynamics of civilian attitudes, the U.S. Army Training and Doctrine Command (TRADOC) Analysis Center (TRAC) developed the Cultural Geography (CG) model to simulate behavioral responses of civilian populations in a conflict eco-system[1].

The CG model is an agent-based model grounded in doctrine and social theory. The model consists of entities (people) interacting with an infrastructure sub-model, interacting with each other through a social network, and responding to specific events. Each entity is defined by a set of demographic dimensions that collectively shape the entity's beliefs, values, interests, stances on issues, and behaviors. Population behaviors are modeled in CG using the Theory of Planned Behavior (TpB) implemented in Bayesian networks[2]. The CG model outputs population stances on critical issues based on various inputs, to include reaction to events, interaction across the social network, and access to essential services.

The CG model is data driven, requiring extensive research and knowledge of the target population's narrative and critical issues. To facilitate the data development process, team 12 tested a proof-of-principle concept for utilizing Tactical Conflict Assessment and Planning Framework (TCAPF) data within the CG model.

The U.S. Agency for International Development (USAID) developed TCAPF in an effort to help civilian and military personnel collect data in unstable areas. The TCAPF questionnaire consists of four open-ended questions[3]:

- Have there been changes in the village population in the last year?
- What are the most important problems facing the village?
- Who do you believe can solve your problems?
- What should be done first to help the village?

TCAPF's straight-forward and effective approach to data collection resulted in acceptance by several U.S. Government organizations in Afghanistan, including the U.S. Army and Marine Corps.

This report describes the team's concept for implementing TCAPF data in the CG model. The team applied the concept using a Pakistan-Afghanistan (PAKAF) case study recently completed by TRAC.

TEAM 12 OBJECTIVE

The primary objective for Team 12 was to explore and implement TCAPF questionnaire data as input to the CG model.

To demonstrate the concept, the team scoped the research to data derived from the second TCAPF question: "What are the most important problems facing the village?" The team selected data from this question because the CG model architecture supports assessment of population stances on critical issues and problems. The benefit of inputting and modeling question #2 data in the CG model is that analysts (and commanders) may gain insights into factors that influence population stances on village problems through experimental designs.

PAKAF CASE STUDY

Team 12 utilized a scenario from the PAKAF Strategic Multi-layered Assessment (SMA) to demonstrate TCAPF data inputted into CG. The PAKAF scenario modeled population stances on three issues from six Helmand province districts in Afghanistan. The three issues under study were security,

[1] Kilkullen, David. Counterinsurgency in Iraq: Theory and Practice. 2007.
[2] Ajzen, Icek. "Constructing a TpB Questionnaire: Conceptual and Methodological Considerations," September, 2002 (Revised January, 2006).
[3] Crnkovich, Mirko. "Tactical Conflict Assessment and Planning Framework (TCAPF) – Stability Operations". USAID presentation dated 26 September 2009.

infrastructure, and governance. For a detailed discussion of the PAKAF scenario, see Hudak et. al, 2010[4].

To support the PAKAF data development process, subject matter experts (SMEs) identified prominent population groups, group beliefs and interests, and events impacting beliefs and interests (such as insurgent attacks or opium eradication operations).

The CG model utilizes Bayesian belief networks to capture the impact of events on beliefs and issue stances. Figure 1 depicts the Bayesian belief network for security implemented in the PAKAF case study. Each entity in the CG model 'possesses' a belief network with unique values in the conditional probability tables that underlie the belief network. Figure 1 depicts beliefs as parent nodes with sample conditional probabilities impacting the population's stance on security.

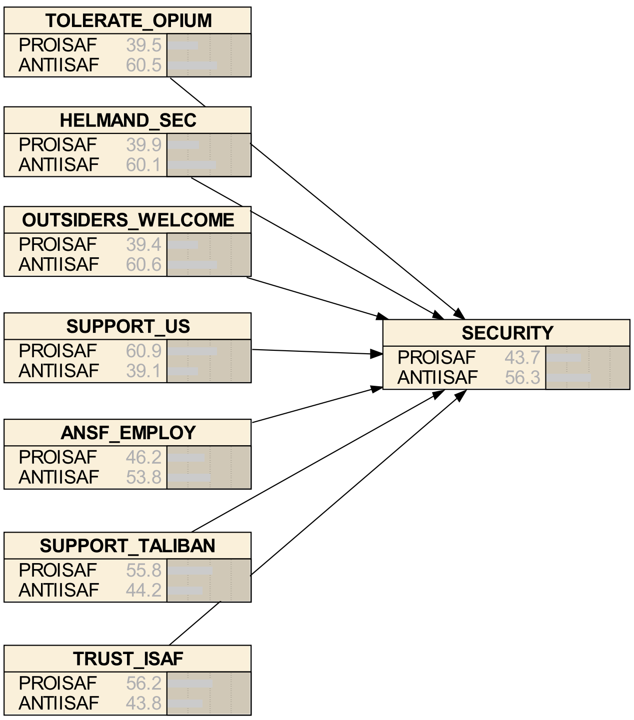


Figure 1. PAKAF Bayesian Belief Network for Security

METHODOLOGY

The team followed the methodology below to input TCAPF data into the CG model:

- Identify and select major problems/issues from TCAPF question #2 for modeling in CG.
- Append selected issues from TCAPF to Bayesian belief networks implemented for the PAKAF case study.
- Map beliefs from the Bayesian belief network to newly appended TCAPF issues/end nodes.
- Develop case files that simulate impact to beliefs (and hence issue stances) resulting from events modeled in the PAKAF case study.

Identify Major Issues from TCAPF Data

The team researched TCAPF data from Helmand province, Afghanistan dated May – September 2009. Respondents to the TCAPF questionnaire resided in multiple districts across Helmand province that generally aligned with the districts modeled in the PAKAF case study.

Results from TCAPF question #2 cited 12 major issues facing the respondents. The team selected four of the 12 problems to model in CG: potable water, irrigation water, education, and health care. Aside from security (which was modeled in the PAKAF study), the four selected issues ranked highest among the respondents.

Append Issues to Bayesian Belief Networks

The team appended the four selected TCAPF issues to Bayesian belief networks developed for the PAKAF case study. The modeling assumption was that beliefs derived from the PAKAF population were sufficiently similar to the beliefs of the TCAPF population. If the beliefs were similar for both populations, then the beliefs utilized for the PAKAF case study could reasonably impact both PAKAF issues and TCAPF issues. Figure 2 illustrates the approach of appending TCAPF issues to the PAKAF Bayesian belief network.

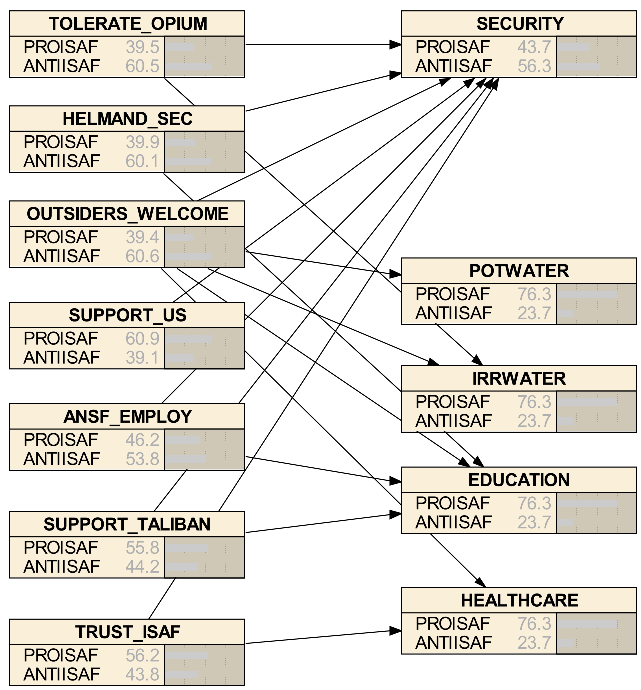


Figure 2. TCAPF Issues Appended to Belief Network

[4] LTC David Hudak, MAJ Francisco Baez, MAJ Steven Jones, Mr Timothy Perkins, Dr Deborah Duong, Mr John Willis, Mr Mark Tanner, Mr Matthew Dearing, Mr Matthew DuPee, Mr Harold Yamauchi, and Mr. Gerald Pearman. "Cultural Modeling Support to Pakistan-Afghanistan (PAKAF) Strategic Multi-layered Assessment (SMA)." Technical Report, 2010.

Map Beliefs to TCAPF Issues

The next step required the team to map beliefs from the PAKAF case study to TCAPF issues appended to the Bayesian belief network. Specifically, the team assessed each belief node to determine whether it would likely impact any of the TCAPF issues. For instance, the belief 'Tolerate Opium' from figure 2 would likely have a relational impact on the population's stance on irrigation water (depicted as 'IRRWATER' in figure 2).

Develop Case Files to Impact Beliefs

The final step in the methodology involved the process of developing case files that impact beliefs in the Bayesian belief network. In the CG model, beliefs may be impacted by events from external actors. For instance, assume the CG model simulates coalition forces conducting opium eradication operations. Following this event, the belief 'Tolerate Opium' would likely be impacted. Assuming 'Tolerate Opium' is mapped to the issue of irrigation water, opium eradication would affect the population's issue stance on irrigation water.

The process of developing case files involved SMEs completing a questionnaire tailored to the events, population groups, beliefs, and issues under study. Specifically, SMEs assessed the impact of each event on each belief from the perspective of each population group. For instance, the PAKAF case study modeled rural and urban population groups. Extending the example above, SMEs might assess that opium eradication impacts the 'Tolerate Opium' belief more for rural dwellers than urban dwellers because rural dwellers are more likely to engage in opium production than urban dwellers. The questionnaire also required SMEs to assess the impact of end node issue stances (to include the four issues from TCAPF) by event and population group.

RESULTS

The team executed an experiment in the CG model involving 14 factors (namely the events modeled in the PAKAF case study) and Bayesian belief networks and case files simulating the TCAPF issues.

The team expects to analyze output from the runs by comparing CG results to TCAPF results with respect to tribal affiliation and occupation. The PAKAF case study modeled population dimensions according to five categories, including tribal affiliation and occupation. TCAPF data also captured respondent demographics by tribal affiliation and occupation. Assuming that TCAPF data is 'ground truth' (or the baseline condition), comparing CG model output against TCAPF output for these demographic groups will provide a measure of validation for the CG model. Results of this analysis will be published in thesis research scheduled for June 2010.

CONCLUSIONS

The team developed and successfully implemented a sound methodology for augmenting a preexisting CG scenario with TCAPF data. Our team's contribution represents a starting point for integrating a popular data collection framework with the CG model.

Recommendations for further research include:

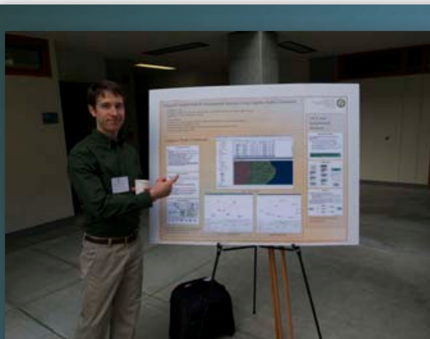
- Improving the CG model to include population migration capability. This capability may enable analysts to model and explore factors impacting TCAPF question #1 data: "Have there been changes in the village population in the last year?"
- Utilizing the CG model to generate simulated TCAPF data. The methodology described in this report facilitates generating TCAPF data from CG. Specifically, the Bayesian belief networks appended with TCAPF issues enable analysts to 'poll' CG entities following model execution to determine issues of greatest interest. This capability would be useful during training exercises and tactical wargames, such as TRAC's 'Irregular Warfare Tactical Wargame.'

BIBLIOGRAPHY

[4] Alt, Jonathan K., Leroy A. 'Jack' Jackson, David Hudak, and Stephen Lieberman, "The Cultural Geography Model: Evaluating the Impact of Tactical Operational Outcomes on a Civilian Population in an Irregular Warfare Environment." Journal of Defense Modeling and Simulation: Applications, Methodology, Technology, 2009.

[5] Department of the Army, Headquarters. "FM 3-24, Counterinsurgency." Washington, DC: Department of the Army, Headquarters, December 2006. 1-14..

Poster Session



Team 13: U.S. Customs and Border Protection (CBP) Analysis Support: Tunnel Detection System

TEAM 13 MEMBERS

Manuel A. Ugarte
Thomas S. Anderson, Ph.D.
U.S. Army TRADOC Analysis Center, US

Thomas Huynh, Ph.D
Gary Langford
Chris Nannini
Thomas McMurtrie
Sarah Wolberg
Naval Postgraduate School, US

Brittlea Brown
Northrop Grumman Corp.

INTRODUCTION

Since 1990, more than 116 cross-border subterranean tunnels have been discovered along the continental US borders, the vast majority between US and Mexico. Tunnels present a low probability, high threat scenario to the United States and are a known means of illicit transportation of drugs, weapons, money and people across the US border. The perpetrators engaged in illicit trafficking are intelligent, tenacious, technologically innovative and they relentlessly seek to continue to expand their profitable enterprise. In today's world, confronted with the realities of terrorism and terroristic objectives, one must also acknowledge that tunnels pose a looming threat to national security. Tunnels are also a persistent military threat. A 2007 operational needs statement (ONS) from US Central Command (CENTCOM) noted that detainees were attempting to tunnel as a means to escape from the internment facilities. Another region with an emerging subterranean threat to US Forces is in Afghanistan with the Karez, or underground aquifers built to move irrigation water from mountains to villages by normal gravity-driven flow. The Karez in Afghanistan (thought to number 6,000) present the Taliban and other insurgents with a means to cache weapons and material, infiltrate and exfiltrate the battlefield and move fighters and supplies. Furthermore, in Egypt, the flow of weapons, ammunition, and other contraband under the Egyptian border has contributed significantly to the ongoing Israeli-Palestinian conflict. Open source estimates place the number of tunnels along the Israel-Gaza border between 300 and 1000. The US Army Corps of Engineers successfully answered the ONS from CENTCOM by developing and implementing the Tunnel Activity Detection System (TADS) (R2TD Implementation Directive).

The objective of the TRADOC Analysis Center - Monterey (TRAC-MTRY) study is to determine the sensor system, to include the TADS, that maximizes the probability and efficiency of detecting existing tunnels and tunnel construction activity on the US border according to geographic location, infrastructure, and historical data. We used a systems engineering and analysis approach to determine the drug traffic organization's (DTO) tunnel infiltration techniques on the Southern US border according to geographic location, infrastructure, and historical data.

This paper provides a high level overview of the systems methodology and describes the initial effort conducted during the IDFW20 to implement an agent- base approach that will enable the analysis of tunnel detection systems. The methodology was comprised of using the functional decomposition of the tunnel threat to create a set of threads and vignettes for analyzing sensor type and allocation with agent-based modeling and simulation. The tunnel threat was based on a threat enterprise model of narcotics trafficking, coupled with global information systems (GIS) data. We evaluated information instances pertinent to tunnel threat behaviors to include historical tunnel locations, urbanization of border towns and tunnel attributes to support a strategy for equipping the border with a persistent tunnel defeat capability.

APPROACH TO THE TUNNEL DETECTION AGENT-BASED MODEL.

This research coalesces exploratory data analysis, analytical models, system engineering and analysis, and modeling and simulation to gain insights into sensor allocation, configuration, placement, and prioritization of sensor field emplacement along the US southern border. We applied operational analysis research methodologies to gain insights into these issues by evaluating the data and technologies available in order to enable a well informed recommendation for the system capabilities required to detect tunnels. The goals of Team #13 were: (a) develop a preliminary design of a prototype Tunnel Detection System asset allocation/trade-off analysis vignette in an agent-based model (ABS) modeling environment, (b) develop a data farming methodology that lends itself to ease of use for analysts and (c) identify and define appropriate system measures.

DESCRIPTION OF SCENARIO

Given a generalized problem statement from the CBP, the physical domain of the problem was described and assessed. The operative effort focuses on building scenarios. Scenario building helps stakeholders make strategic decision to adapt to their several possible futures. The focus is to identify the main driving forces and areas of uncertainty.

Based on this physical domain, the threats, CBP missions, and constraints of the system were used to develop a series of scenarios (a narrative description of possible future events) each exemplified by vignettes. The vignettes (operational responses to the possible future events) were used to identify the system's actor-driven use cases. The vignettes served as the baseline for diagramming various behaviors using the systems modeling language (SysML). The use case diagrams enable initial scenario development to begin through implementation of the system's activity diagram, sequence diagrams, and functional and process decompositions, which aid in the construction of the system's concept of operations. Next, the design parameters and factor levels affecting CBP's sensor feedback and TTPs were identified. The model will be translated into an executable agent-based model, where an appropriate experimental simulation design will be applied. The results of the simulation will be captured and analyzed in order to determine an appropriate sensor outlay and which, if any, specific tunnel interdiction TTPs should be implemented by CBP.

The initial, overarching detection system use case diagram is depicted in Figure 1. The overview CBP detection system of systems includes various human elements, networked sensor systems, and key data links between the internal system and external data sources.

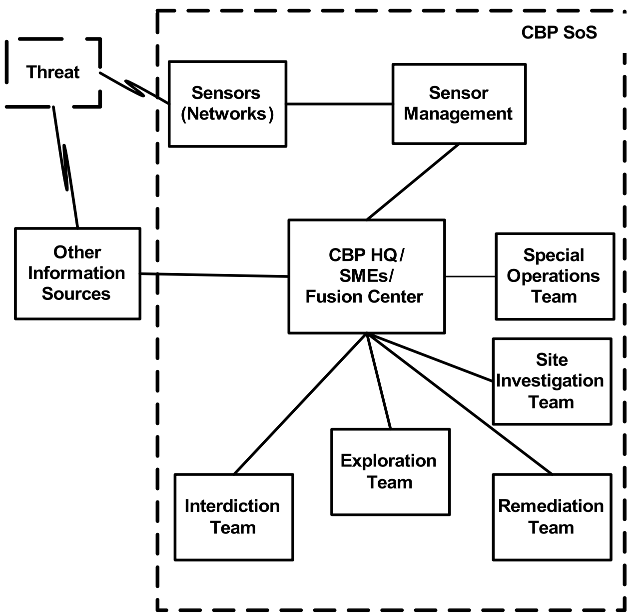


Figure 1. CBP Baseline Use Case.

After development of the baseline use case, the team applied a threat vignette (in this case, prosecution of a detected tunnel) to develop a sequence diagram depicting the process activities and functional actions of each of the relevant elements. The sequence allowed for the development of an initial system concept of operations. This CONOPS is summarized in the subsequent paragraph below.

A specific signature threshold is picked up by an array of underground sensors. The observable measurements are transmitted to the sensor management station as well as the Subject Matter Expert (SME) / Headquarters (HQ). The SME may request more information from the sensors and the management station. Concurrently, the SME may also receive external information from other sources. The sensors will respond to these requests and transmit additional parameters and observations to the SME. This data will include threat signature threshold, range, and position.

These observations allow the SME to further integrate and analyze the data and information. Once the SME has evaluated and interpreted the data to indicate that a tunnel exists, he/she will notify the Site Investigation Team to investigate the situation further. The Site Investigation Team is made up of four members. Their job will be to conduct an area site survey of the specified location. Also, they will estimate the cost of the tunnel discovery, to include determining if and where a tunnel or cavity exits. Once a tunnel is located, they will transmit their results to the SME/HQ in addition to notifying the Interdiction Team that is comprised of a Special Operations Team and Exploration Team. The Interdiction Team will investigate the specified location and drill a series of holes using a 1" – 4" auger drill bit to verify the suspected location of the tunnel. A larger bore will be implemented for ingress of robots or Special Operations teams for analysis and exploitation. They will determine the tunnel's characteristics and type which will subsequently be transmitted to the SME. The SME will direct to initiate tunnel closure actions by the Remediation Team. The team will fill the tunnel with an appropriate material (e.g. slurry substance or concrete) and will submit a tunnel closure report when complete.

In the case that the Site Investigative Team concludes that no tunnel is found, they will notify SME/HQ that no tunnel is found and flag the area and report a false alarm. The SME/HQ will develop a historical document trail for the area and note the region as a "yellow" or potential hot region for continued or increased monitoring activity.

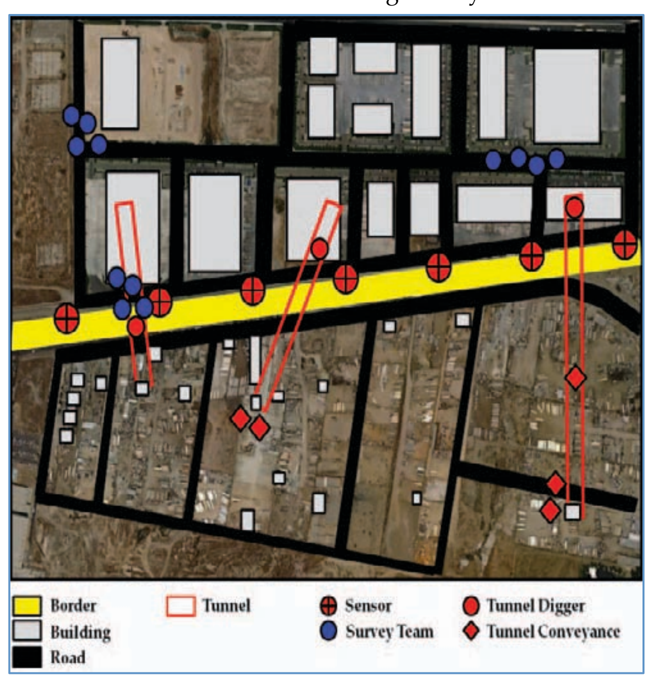


Figure 2. Agent-based Model Scenario Concept Sketch.

Based upon these CONOPs, initial efforts were made at modeling the scenario in Pythagoras (an agent-based model that is suited for data farming, execution of large numbers of repetitions of parametric runs to identify behaviors overlaid on a dynamic landscape). A concept sketch and screenshot of the base case scenario implemented in Pythagoras are shown in Figures 2 and 3, respectively.

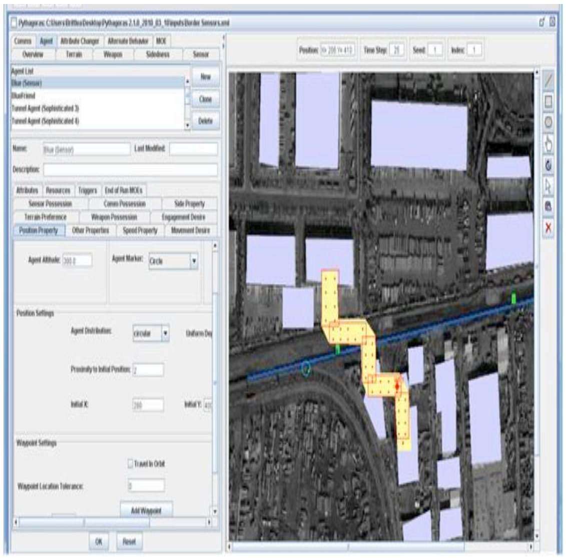


Figure 3. Agent-based Approach in Pythagoras.

Design of Experiments

The design of experiment (DOE) methodology will help us to gain insights on sensor system behavior and its interaction with key factors. Our goal is to gain such insights from factors that influence the disruption of tunnel construction and use. We will use Pythagoras, developed by Northrop Grumman, to represent tunnel detection and interdiction assets along a cross-border region of interest. During model development, we created notional factors (Table 1) that define both the characteristics and performance capabilities of the CBP sensor management system and the tunneling activities that it was designed to defeat.

Our scenario consists of a region along the United States/Mexico border with representative road and building infrastructure as previously depicted in Figures 2 and 3. We will develop agents to represent the CBP system of systems (SoS) and threat as represented in Figure 1. Agents will include team members, vehicles, and ground sensors that encompass the SME/HQ Fusion Center, Site Investigation, Interdiction, Special Operations, Exploration, and Remediation teams, and associated sensor networks. The threat agents will consist of tunneling construction entities (digging and conveyance).

Our specific design will capture the factors relevant to the CBP SoS. Such factors include the type and number of sensors employed, sensor performance, configuration and field emplacement characteristics, the number and behavior of members in the CBP teams, as well as the characteristics of tunneling entities.

We anticipate that numerous factors consisting of multiple levels will be required to model the scenario. Employing a full factorial experiment with the 23 factors

Scenario Factors	Factor Name	Units	DOE Ranges	
			Low	High
Seismic Sensor				
Number of Sensors	NumSens	#	5	15
Probability of Detection	SensDetect	N/A	0.60	1.00
Range of Detection	SensRange	meters (m)	100	300
Target Location Error	SensTLE	N/A	0.60	1.00
Emplacement Distance	SensDistance	meters (m)	10	110
Survey Team				
Number of Survey Teams	NumSurvey	#	1	3
Movement Speed	SurveySpeed	meters/minute	60	100
Probability of Detection	SurveyDetect	N/A	0.60	1.00
Range of Detection	SurveyRange	meters	50	250
Target Location Error	SurveyTLE	N/A	0.60	1.00
Time for Survey	SurveyTime	minutes	30	90
Transition Time	SurveyTrans	minutes	30	90
Tunnel Digger				
Number of Tunnels	NumTun	#	2	6
Number of Diggers (per tunnel)	NumDig	#	1	3
Dig Rate	DigRate	meters/hour	1	5
Probability of Counter-Detection	DigDetect	N/A	0.60	1.00
Dig Delay Time	DigDelay	minutes	30	90
Tunnel Transport				
Number of Transports (per tunnel)	NumTrans	#	1	3
Transport Speed	TransSpeed	meters/minute	20	60
Probability of Counter-Detection	TransDetect	N/A	0.60	1.00
Headquarters				
Number of Headquarters	NumHQ	#	1	3
Communication Rate	CommRate	message/hour	4	60
Communication Error	CommError	N/A	0.60	1.00

Table 1. Notional Factors Representing the CBP Sensor Management System and Tunneling Activities.

presented in Table 1, each set at two levels (low & high), would result in 323 design points. A pairwise projection for the first four factors of such a design is displayed in Figure 4. Fortunately, there are more efficient alternatives to using a full factorial design.

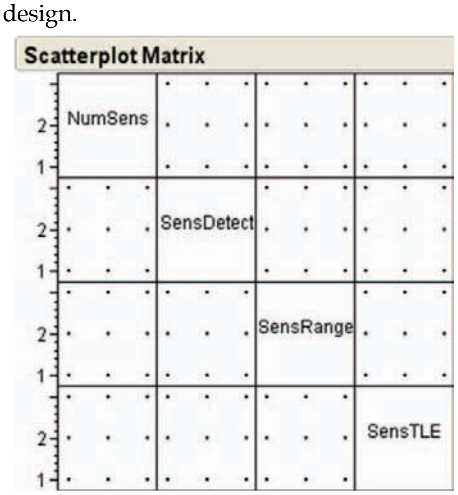


Figure 4. Scatterplot Matrix of a Factorial Design for Four Factors.

The space-filling feature of Nearly Orthogonal Latin Hypercube (NOLH) enables efficient exploration of the solution space represented by measures of effectiveness (MOEs). An NOLH experimental design will facilitate a comprehensive analysis of potential explanatory variables within our model. Rather than being restricted to two or three levels, the analyst can create a design that uses multiple levels or even a continuous range of values for each factor.

In addition to the space-filling property, orthogonal designs have no linear relationship between the regressors. The NOLH technique minimizes the correlation between factor columns, creating a nearly orthogonal design matrix. We can examine the off-diagonal elements within the correlation matrix of the design in order to measure the level of orthogonality.

Using the NOLH designs spreadsheet tool developed by Professor Susan M. Sanchez at the Naval Postgraduate School, we generated 257 design points consisting of 23 factors. Our preliminary design consists of a 257 x 23 matrix with the largest correlation of 0.089.

This design is displayed in Figure 5 which represents the space-filling, two-dimensional projections for the first four notional factors.

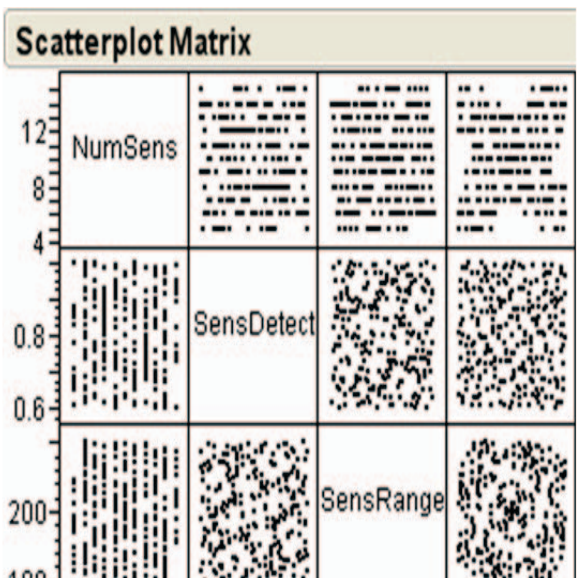


Figure 5. Scatterplot Matrix of NOLH Design for Four Factors.

SUMMARY

Analysis of Base Case

The outputs of interest consist of the possible sensor system allocations for a given CBP concept of operation. The outputs are highly multivariate, consisting of all sensor systems positions, dispositions, and strengths as well as the perception of tunnel location. That is, the output is a Common Operational Picture (COP) comprised of the adversary and/or CBP tunnel interdiction assets. The MOEs that may be derived from ABM include; time, assets and money allocation. In this manner we determine the constraints that will drive a future response.

Further Work

The challenges to tunnel defeat are attributed to the vast lengths of the border, varied geologies that degrade persistent monitoring and geophysical resolution of subsurface anomalies, as well as the unknown quantity of existing tunnels. The tunnel denial and defeat mission must be addressed in a comprehensive manner in order to be successfully implemented. This means embracing a heterogeneous system of systems approach that allows both passive and active sensing, command and control, and robotics for post discovery interrogation. Before committing to a particular technical solution, a cost-benefit study and an analysis of systems study that takes into account operational performance associated with GIS locations, risk assessment associated with GIS location, man power requirements, projected training, maintenance costs, and sustainability factors will provide the necessary oversight and will demonstrate due diligence in specifying a system solution.

Within the context of systems engineering and analysis, agent-based modeling for tunnel detection systems shows great promise as a method to enhance the evaluation of multiple detection systems under various conditions. In addition, further work may include developing and testing a DOE capability and designing a user interface appropriate to the skill level of the operator, the required response times, the resource allocations, and the operational environment. Improved automation of output analysis is also desirable, especially formatting the output reports to be more amenable for use by statistical packages for data analysis, evaluation, and assessment.

REFERENCES

[1] US NORTHCOM's Rapid Reaction Tunnel Detection (R2TD) Implementation Directive, FY10-12 Joint Capability Technology Demonstration, 15 March 2010.

[2] Huynh, T.V. and J.S. Osmundson. An integrated Systems Engineering Methodology for Analyzing Systems of Systems Architectures.

[3] Huynh, T.V. and J.S. Osmundson. A Systems Engineering Methodology for Analyzing Systems of Systems Using the Systems Modeling Language (SysML).

[4] Cioppa, T. M. and T. W. Lucas. 2006. Efficient nearly orthogonal and space-filling Latin hypercubes. Technometrics.

[5] Sanchez, S. M. 2006. Work smarter, not harder: Guidelines for designing simulation experiments. Proceedings of the 2006 Winter Simulation Conference, eds. L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto, forthcoming, Piscataway, NJ:.Institute of Electrical and Electronic Engineers.

Team 14: Anti-Piracy and Terror Reduction

Simulating Pirate Behavior to Exploit Environmental Information

TEAM 14 MEMBERS

Leslie Esher, LT USN
Stacey Hall, LT USN
Jim Hansen, PhD
Eva Regnier, PhD
Paul Sanchez, PhD
Dashi Singham, PhD Candidate

Problem

Due to the increase in pirate activity off the coast of Somalia, the United States military and the combined forces of the world's navies are partnering together to defeat these violent extremists. Piracy has threatened maritime safety and cost commercial shipping billions of dollars paid in ransom monies. The Gulf of Aden and the Horn of Africa that were once safe to transit are no longer, and for this reason, President Obama has issued an executive order to defeat terrorism in the form of piracy. The Commander of the U.S. Naval Forces Central Command (CENTCOM), U.S. Fifth Fleet, Combined Maritime Forces (CMF), is responsible for the safety, stability, peace, and vital interests of the United States for 2.5 million square miles of water. For the contents of this paper, the region of geographical concentration has been on the area that has been the most prevalent to pirate attacks, Somalia. Combined Task Force 151 (CTF 151) is a multi-national task force that is response for 1.1 million square miles of water in the Gulf of Aden and off the coast of Somalia.

Pirates in this area generally operate from small boats (skiffs) that have limited survivability at sea in severe weather conditions; this paper will refer to these conditions as METOC, (Meteorology and Oceanography) conditions. High sea state and/or wind speeds make it difficult or nearly impossible for pirates to attempt to board commercial vessels. The analysis of this paper is to provide insight into what parameters are most influential in contributing to and limiting pirate behavior.

In response to the piracy problem, the U.S. Naval Oceanographic Office (NAVOCEANO) at Stennis Space Center has been providing a forecasting product, called the Piracy Performance Surface (PPS) that uses forecasts of winds and seas to map the locations that are most conducive to pirate activity, and incorporates information on confirmed pirate activity in the form of an attack, an attempted attack, or suspicious activity. The existing product was developed rapidly to provide support to the operators. NAVOCEANO is working to improve the model of the relationship between METOC and pirate activity, and to improve the way the pirate threat is updated when confirmed piracy activity is observed.

The overarching research question is: How can the N2/N6 (Director for Information dominance that comprises information, intelligence, command, and control) contribute decision-critical information to the operators who are protecting commercial shipping traffic.

By September 2010, a new simulation-based engine will be implemented to produce the PPSNext. The simulation is based on a model of pirate behavior (hereafter, CONOPS, Concept of Operations), combined with forecasts for METOC conditions and intelligence on certain parameters of pirate behavior, such as whether they operate from land or sea bases (mother ships) and the number and locations of these bases.

The goal of Team 14 was to provide insight on which parameters describing pirate CONOPS were the most important drivers of the map reflecting relative pirate threat and which have the strongest interaction with METOC variables. These results would indicate which factors in pirate CONOPS are most important to include in the model and which parameters should receive most intelligence resources.

Simulated Pirates and Environment

In the model of pirate CONOPS, the basic strategy is to depart from a base – either a land-based camp or a sea-based mother ship – typically a Boston Whaler that has longer longevity and life expectancy at sea, with a handful of pirates with a few days' supplies. The skiff motors to its pre-determined location (latitude and longitude). As illustrated in Figure 1, the skiff then drifts with the winds and currents

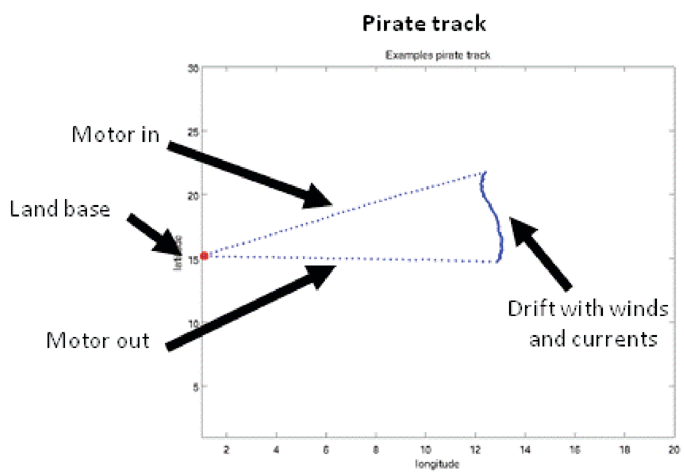


Figure 1: Example of a pirate skiff trajectory

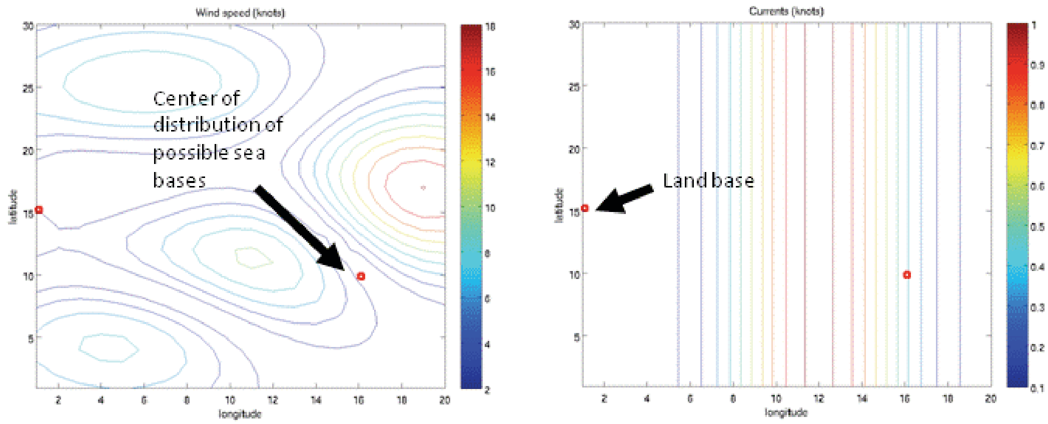


Figure 2: Simulated winds and currents.

until the pirates run out of supplies, at which point the skiff motors back to its land or sea base.

Winds, waves and currents affect the pirates. In their drift phase, their motion is determined by currents and winds. In addition, one of the factors whose impact we are evaluating is whether pirates use weather forecast knowledge to plan and implement an attack. In the current implementation, if the pirates have forecast knowledge, is it assumed that their information is "perfect". If they have forecast knowledge, they do not go to locations with unacceptable weather-as determined by wind and wave thresholds. If they do not have forecast knowledge and encounter unacceptable weather, they return to their base location.

In its operational implementation, the METOC conditions will be the result of a coupled atmospheric-oceanic model. In the version used experimentally for IDFW-20, notional winds, seas, and currents were used (shown in Figure 2) that changed over the course of the 72-hour simulation, but otherwise did not vary as a function of simulation trial. The pirates operate in a 20×30-cell grid, with each cell 10 km on a side.

Output Statistics

Considerations

Perhaps the biggest challenge this group faced was how to summarize the simulation's output. Although there is a limited database of historical pirate attacks, it has not been possible so far to recreate the METOC conditions corresponding to the period of known pirate activity against which to verify the model. Therefore, for any experiments conducted during the IDFW, there is no ground truth against which to compare results. In addition, even if we consider only the relative density of pirate activity across the simulated area and summarize pirate activity in 12-hour periods, each simulation produces a pirate density in each of 600 cells at each of six time periods (See Figure 3). Each simulation must be summarized and compared usefully with the results for other design points, to identify the variables that are most influential and most related with METOC conditions.

As described below, we undertook an experiment with 33 design points (simulations), and therefore 528 pairs whose similarity or difference might be measured. Within each simulation, differences across the six time periods would reflect sensitivity to METOC conditions (which changed over

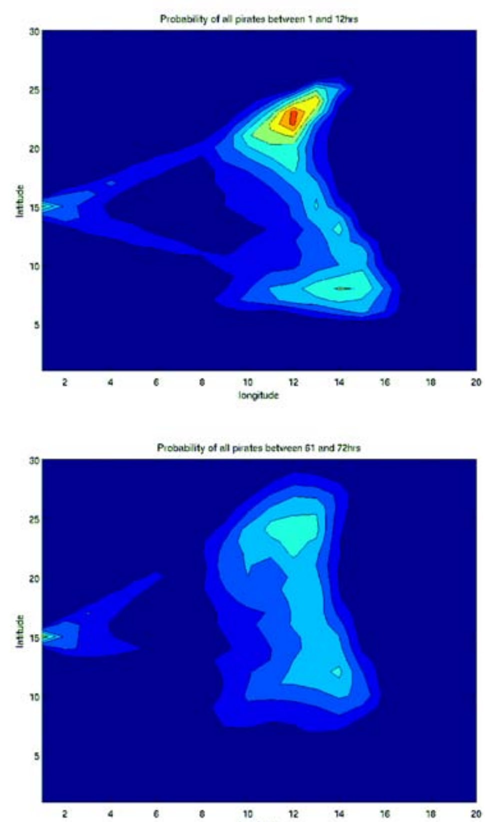


Figure 3: Sample PPSNext simulation output.

the course of the 72 simulated hours) and interactions with METOC conditions. Comparing the six pirate density plots would result in 15 pairwise comparisons.

Potential summary statistics

We considered summary statistics for comparing two maps of pirate density (whether they were from the same time period, and different design points, or the same design point, but different time periods). The following Y-variables were used in our experiment:

- The maximum root mean square cell-by-cell difference (RMSE) between each design point and all other design points. RMSE especially penalizes large errors.
- The cell-wise maximum difference (MaxDiff) in pirate density between each design point and all other design points
- For each trial, the largest RMSE between a 12hr pirate density and the 72-hour averaged pirate density map (this response variable is called ΔRMSE, and analogous measures are ΔMaxDiff and Δ50th prctile)
- The mean across time-periods of the area that bounds 50% of the pirate density (50th prctile).
- Same as each of the above, but for smoothed distributions (indicated by S-prefix).

Other summary statistics that we considered, and which might be applicable in future work include:

- Cell-by-cell differences in mean relative entropy.
- Location: minimum distance between two modes (or sum of minimum 2 or 3 distances).
- Decision-related
  - How much of the total pirate density can be captured within miles of optimally deployed search assets?
  - How big would circular covering disk have to be to capture 75% of the pirate density?
  - Sum of differences over larger ("coarse-grained") cells that might be defined according to the size of an area searchable by Task Force 151 assets within a given time?
- Other:
  - Fourier transform
  - Max eigenvector

Experimental Design

Because the current implementation of the simulation is in Matlab, and therefore we expected each trial to take an hour or two to run, we knew we would be limited in the number of design points. To capture the effects of all variables and interactions among them, we used a Nearly Orthogonal Latin Hypercube (NOLH) design (Cioppa & Lucas, 2007), downloaded from SEED center website. We restricted ourselves to experiment with eleven input variables (X-variables), which allowed us to use an experimental design with 33 design points, which we anticipated was few enough that we could complete the runs overnight. The variables and their maximum and minimum values are shown in Table 1 below.

Variable	Minimum	Maximum
Simulated pirates per day	200	1200
Mission length (hours), *Length*	72	120
Pirate groups	3	7
Total number of land and sea bases*	3	7
Proportion of bases that are sea bases*	0.25	0.5
Known base locations (Yes/No)	No	Yes
Transit speed (kts), *Speed*	8	12
Pirates' wind threshold (kts, *Wind*	10	20
Pirates' wave threshold (ft),*Wave*	3	10
Probability that pirates use forecasts	0	1
Wind drift factor, *Drift*	0.1	0.75
*Used to calculate number of land bases (*Camps*) and sea bases (*Sea Bases*)		

Table 1: Input variables (X) used in the experiment

RMSE	S-RMSE	MaxDiff	ΔRMSE	S-ΔRMSE	ΔMaxDiff	ΔS-MaxDiff	50th-prctile	S-50th prctile
0.77	0.68	0.38	0.55	0.77	0.66	0.84	0.73	0.71
Length × Wind	*Length*	*Speed × Drift*	*Wind*	*Wave*	*Wave*	*Wave*	*Length*	*Length*
Wave^2	*Wind*	*Drift*	*Wave*	*Wind*	*Wind*	*Wave*^{2}	*Wind*	*Wind*
Wave	*Wave*	*Speed*		*Camps × Wave*	*Sea Bases*	*Camps*	*Sea Bases*	*Speed*
Drift	*Length × Wind*			*Camps*	*Sea Bases × Wave*	*Wind*	*Speed*	*Sea Bases*
Wind	*Wave*^{2}					*Camps × Wind*	*Sea Bases × Speed*	
Camps × Drift								
Camps								
Length								

Table 2: For each output measure, adjusted R2 and input variables included in the fitted regression model, ordered from most to least influential.

Results

For each Y-variable, we used JMP statistical software to fit a regression model to a set of 75 potential predictors, i.e. the variables shown in Table 1, their squares, and all first-order (pairwise) interaction terms. JMP performed stepwise regression, allowing variables to enter and leave the model based on their significance (p-value).

Table 2 shows the X-variables that were included in the model, as well as the R2 and adjusted R2 for each model. Wave threshold and wind threshold proved to be the most significant in the current Matlab implementation of the PPSNext model, and at least one of these two variables appeared in the model for every Y-variable.

The absence of some X-variables is very interesting. For example, these results seem to indicate that it is not important for intelligence to learn whether pirates can acquire and use METOC forecasts, nor would it change the PPSNext if they acquired that capability.

The Δ-prefixed Y-variables measure differences within a single simulation (design point) over the 72-hour simulated time period, rather than differences relative to the other design points. Therefore, the X-variables that are most related to the Δ Y-variables can be interpreted as those that have the largest interaction with METOC conditions. Both wind and wave thresholds appear in the model for every Δ Y-variable, indicating that (not unexpectedly) wind and wave thresholds interact strongly with METOC conditions in determining the spatial distribution of pirate activity. The wind drift factor and mission length do not appear in any of these models, however, indicating that they do not interact strongly with METOC conditions.

X-variables that might be estimated using intelligence also appear to drive the results, in particular mission length. The interaction between mission length and wind threshold in two of the models is interesting. The number of sea bases or camps – which are highly related, as the number of sea bases is a fraction of the total number of bases – appear in many of the models, indicating that it would be valuable to have good estimates of the number of bases.

The results do not provide clear guidance as to which of the output measures are more useful. In addition, smoothing does not have a consistent effect on the significance of the results. For some measures, the smoothed output model achieves greater R2 than the raw value and for some measures the opposite. The smoothed MaxDiff did not produce any X-variables that were significant at least at the 0.01 level, and therefore its model is not shown in Table 2.

Future Work

Near-term future work on this project (in the next year) includes running similar experiments using the operational code, which will include environmental and navigational conditions for specific, real area of operations, in particular the area off the HOA plus the Gulf of Aden. We will seek to confirm the qualitative results of the experiments conducted during IDFW-20, to identify which aspects of pirate CONOPS are most critical in interaction with METOC conditions and METOC uncertainty.

Another major component of future work in the next year is the possibility of building an agent-based model that will be able to represent other factors that we know to be important to the problem of detecting and protecting against the pirate threat. In particular, we would like to add agents that represent commercial shipping, the searchers and neutral vessels. We spent some of our time researching environments for implementing an agent-based piracy model and the key features that we would like to see included.

Donna Middleton (Northrop Grumman) gave us a demonstration of Pythagoras, including a simulation she created to capture the effects of currents and waves. Pythagoras has the flexibility to model pirate behaviors such as seasickness and incorporate behavioral habits where pirates run out of cot, food, or water so they return to their origin at different times for a single time step simulation. Another great asset for agent based modeling is the ability to run a multitude of simulations quickly. The downfall of using this agent-based model is the inability to model METOC as fluid dynamics since weather conditions change with each time step. METOC would be static while the agents would be dynamic. Although this feature is not represented in the current pirate simulation, it would be nice to allow agents to have imperfect information about METOC conditions, representing a forecast.

Mary McDonald also visited our team to discuss the applicability of MANA to this problem. We took some time to analyze the pros and cons between each agent-based model. MANA did not have the model flexibility that we so desired with modeling pirate agents and it too has the inability to model changing METOC conditions. Abel (2009) used MANA productively to model frigate defense effectiveness against pirate activity because MANA enabled him to model quadrant dimensionality of the frigate in the form of port, starboard, fore, and aft.

References

[1] Abel, H. Frigate Defense Effectiveness in Asymmetrical Green Water Engagements . Master's thesis, Naval Postgraduate School, Monterey, CA>

[2] Cioppa, T.M. & Lucas, T.W. (2007) Efficient Nearly Orthogonal and Space-Filling Latin Hypercubes. Technometrics 49(1):45-55.

[3] Murphy, M.N. (2009) Small Boats Weak States Dirty Money: Piracy and Maritime Terrorism in the Modern World. Columbia University Press, New York. 539 pp.

Team 16: Logistics Battle Command Model

TEAM 16 MEMBERS

MAJ Francisco Baez
U.S. Army Training and Doctrine Command Analysis Center in Monterey, USA

Mr. Jonathan Shockley
U.S. Army Training and Doctrine Command Analysis Center in Fort Lee, USA

CDR Carston Sultz, GER NAVY
Naval Postgraduate School

Overview

This working group explored the use of the newly developed Logistics Battle Command (LBC) model prototype Graphical User Interface (GUI). Team 16 focused on the development and implementation of a logistics scenario designed to assess the operational impact of different strategies for the management and allocation of transportation assets. Particularly, the specific objectives of Team 16 sessions during IDFW 20 included:

- Implement the scenario in LBC.
- Use visualization features of the GUI to build and analyze the scenario.
- Use the analysis features of the GUI to understand the output results.
- Report the findings.

LBC MODEL

The LBC model is a low-resolution, object oriented, stochastic, discrete event model programmed in Java that incorporates Simkit as the simulation engine. LBC serves as a stand-alone analysis tool or as a dynamic logistics module that can be fully integrated into an existing combat model. LBC functionality includes planning and decision support features to enable a simulated sustainment decision maker to monitor the logistics common operating picture, forecast demand for most classes of supply, and initiate and adjust missions to distribute supplies and perform sustainment functions. LBC uses network architectures to represent the distribution pipeline to summon sustainment planning and execution representing the end-to-end flow of resources from supplier to point of consumption. LBC accomplishes this overall representation through three layers of network representation: the transportation, communications, and planning networks.

The bottom layer is the transportation network. This layer links the LBC model to the physical area of operations representing the geographical distribution of supplies. Algorithms within LBC generate missions including determining the best methods and routes for transporting supplies to the end user while accounting for changing battlefield conditions.

The middle layer is the communications network. This layer represents an arbitrarily complex communications network of the distribution system linking leaders and Soldiers to all applicable stakeholders including the logistics common operating picture. It carries the data of the distribution system information network and links the planning and transportation layers in the LBC.

The top layer is the planning network. This layer represents the data of the distribution system information network. The planning network in LBC allows for monitoring any deviations between the sustainment execution and the sustainment plan, and also allows for sustainment re-planning. LBC uses a task network to link the sustainment planning to execution.

The current LBC version is the result of substantial revisions and expansion by the U.S. Army Training and Doctrine Command Analysis Center to improve the functionality and usability of the model as an analysis tool. To bridge the gap between scripted logistics planning and dynamic re-planning, research is being pursued and several modifications are currently being implemented to enhance LBC's functionality, namely, dynamic movement of units, advanced forecasting, rule-based decision making, and dynamic re-planning. In addition, a prototype GUI is in the testing and validation phases. The GUI (which is considered more of an analyst interface) provides the user with the ability to quickly construct, visualize, and analyze scenarios. The end result of the aforementioned functionality enhancements will be a simulation model to represent a simulated decision-maker to monitor the logistics common operating picture and the capability to select and execute corrective actions in response to operational tempo and enemy activities that require deviation from a predetermined logistics support plan.

SCENARIO

The scenario implemented during IDFW 20, was a basic logistics sustainment support operation in a peace-keeping operation combined with humanitarian assistance. The scenario was set up to regularly push supplies from points of debarkation to consumers in an explicit area of operation. For this scenario the team developed a task network to model the activities involved in moving supplies from source to destination. The model reordered and measured the time between specified activities.

CONCLUSIONS

The work accomplished throughout IDFW 20 was valuable. Despite the several challenges encountered during the week, Team 16 was able to implement the scenario using the prototype GUI; however, the team was unable to conduct any analysis. Nevertheless, the Team 16 participants gained experience using the prototype GUI and several model improvements were identified and implemented during the week. But above all, the prototype GUI proved to be extremely useful for developing, implementing, and debugging a scenario as well as very practical for the representation and analysis of complex, communications, transportation, and planning networks.

International Data Farming Workshop 21

When: 19-24 September 2010

Where: Lisbon, Portugal

Hotel: TIVOLI ORIENTE; Single: 105 euros; Double: 115 euros. This rate includes breakfast and is guaranteed at this price until 22 August. For Hotel Registration, book directly mentioning code IDFW-AM Set2010 and please inform fjv.freire@gmail.com. To make a reservation at the Tivoli Oriente, email Carmen Berimbau at comercial3.hto@tivolihotels.com with the dates you would like to reserve a room for and mention IDFW 21.

Conference Fee (still to be determined, but approximately 400 euros) to cover meeting rooms, buffet lunches, coffee breaks, opening dinner, etc. Any questions, please contact: Colonel FERNANDO FREIRE fjv.freire@gmail.com or LTColonel JOAO ROCHA: j.rocha@oniduo.pt

Registration Form: will be available at https://harvest.nps.edu soon. For questions please contact Gary Horne at gehorne@nps.edu or 831-233-4905.

IDFW 21 Tentative Agenda

- **Sunday, September 19:** Opening reception and dinner
- **Monday, September 20:** Opening briefs and team poster sessions in the morning, then begin work in teams
- **Tuesday - Thursday, September 21 - 23:** Work in teams (optional plenary sessions in the mornings)
- **Friday, September 24:** Outbriefs and Closing Ceremony in the afternoon

Call for Team Leaders / Plenary Speakers:

Please email gehorne@nps.edu if you want to lead a team or present a plenary briefing.

Theme: Discovery

International Data Farming Workshop 21

September 19-24, 2010

Monterey, Portugal

Scythe - Proceedings and Bulletin of the International Data Farming Community
Seed Center for Data Farming
273 Glasgow Hall ~ Naval Postgraduate School ~ Monterey, CA 93950

Issue 8 - Workshop 20