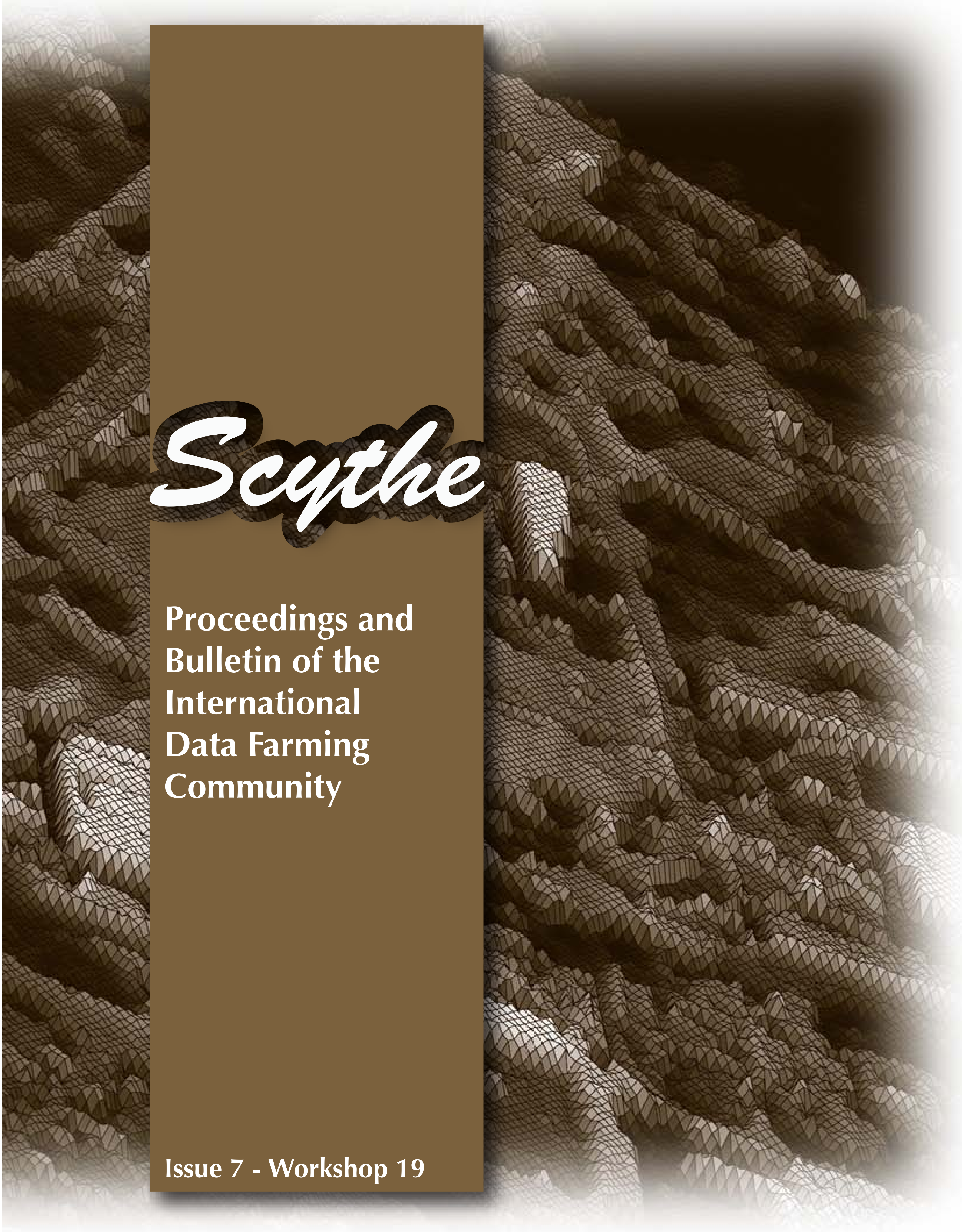


Scythe

Proceedings and Bulletin of the International Data Farming Community

Issue 7 - Workshop 19

Proceedings and Bulletin of the International Data Farming Community

Table of Contents

IDFW 19: "Mana" 1

Team 1: Counter-Piracy Study using MANA........................ 2

Team 2: Impact of Logistics on Readiness and Life Cycle Cost 5

Team 3: Analyzing Selected Questions in a Refugee Camp Scenario Using PAX3D 12

IDFW 19 Plenary Sessions and Focus Groups.................... 16

Team 4: Evaluation of Electro-Optical Sensor Systems in Network Centric Operations Using ABSEM 0.3 17

Team 5: Force Protection with ITSim (II): Base Protection Against Ballistic Weapons......... 20

Team 6: Utility of Distillation Modeling for Countering IEDs........................ 27

Team 7: Social Simulations and Survey Data Focus Group 31

Team 8: Data Farming in Support of NATO 32

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.

Please let us know what you think of this *seventh* prototypical issue. Articles, ideas for articles and material, and any commentary are always appreciated.

International Data Farming Community Workshop 18 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 7, Workshop 19 Publication date: February 2010

Bulletin Editors

Gary Horne:.....................................gehorne@nps.edu
Andrea Hernandez: ....abhernan@nps.edu
Ted Meyer: ....temeyer@nps.edu

IDFW 19: "Mana"

by Gary Horne
Naval Postgraduate School

International Data Farming Workshop 19 was held in Auckland from November 1st through 6th, 2009. This workshop was the third we have held in New Zealand and as usual, it was a magnificent venue! Our theme was "Mana." In the Maori culture, having mana means to have influence, authority, effectiveness, power, usefulness, and prestige. And our goal, as usual, was to use our mana as well as our data farming methods to explore our important questions.

On behalf of our host Dr. Michael Lauren from the Defense Technology Agency of New Zealand, I would like to express our thanks to the team leaders, the plenary speakers and all of the participants in IDFW 19. And, of course, a large Thank You to Michael and all of the folks from New Zealand who worked to make IDFW 19 a world-class event!

This issue, our seventh, 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 back in Monterey, California, USA for our next workshop, International Data Farming Workshop 20. I would like to invite you to participate, starting with the pre-workshop dinner on March 21st, 2010. The workshop will be held from March 22nd through March 25th, ending on Thursday this time. Our theme for IDFW 20 is... "Enrichment." We hope to see you there!

Gary Horne

Team 1: Counter-Piracy Study using MANA

TEAM 1 MEMBERS

Mark Anderson (lead)
Dr Michael Lauren
DTA, New Zealand

LT Brock Symmons
RNZN, New Zealand

Xu Yong Liang
Tiang Si Jia
DSO, Singapore

Lau Joon Hong
DSTA, Singapore

Dr James Decraene
NTU, Singapore

Marc West
DSTO, Australia

LT Travis Gill
NPS, United States

Bernt Akesson
PVTT, Finland

INTRODUCTION

This study investigated a counter-piracy scenario using MANA V in conjunction with the data farming process to explore options for defending a large vessel, relying principally on non-lethal deterrents.

SCENARIO DESCRIPTION

A generic large unarmed vessel (Blue) is transiting through a wide area of water at an economical cruise speed. An agent representing a cluster of pirates (Red) then closes in on it. Once Blue determines their intent, it attempts to manoeuvre away and if required, employs a non-lethal weapon (NLW) to disrupt Red.

The critical assumptions in this scenario concern the rules of engagement. Firstly, it is assumed that Blue would not employ evasive manoeuvres or a non-lethal weapon until the intent of Red was established. Secondly, it is assumed that Red would not attack if the targeted vessel was being escorted by a warship.

KEY MODEL PARAMETERS

The Blue force consisted of a single vessel with a simple cookie-cutter sensor, a non-lethal weapon and a reasonably high level of inertia (resistance to acceleration).

Factor (units)	MANA Value
Cruise Speed (kt)	7
Manoeuvre Speed (kt)	15
Determination of intent range (m)	2,000
NLW range (m)	100
NLW hit probability	0.2

Table 1: Blue Force base parameters

The non lethal weapon was modelled using the refuel parameter in MANA V, noting that Red must first be classified before this can be employed.

The Red force utilises fast highly-manoeuvrable attack craft, such as skiffs with outboard motors. The following table summarises the Red MANA base parameters:

Factor (units)	MANA Value
Speed (kt)	25
NLW Speed (kt)	5
Inertia (unitless)	0.1
Fuel (unitless)	1000

Table 2: Red Force base parameters

When influenced by Blue's non-lethal weapon, Red was sent into a trigger state in which its speed was reduced to 5 knots.

It was assumed that Blue would surrender if Red closed to within 10 m – that is, when Red was close enough to board them.

Fuel was used to model Red motivation, with Blue's NLW degrading it with increasing engagement time. When the fuel level reached zero, Red was deemed to have run out of motivation. This occurred after 15 minutes of engagement time.

EXCURSIONS

Team 1 started by fine-tuning a baseline scenario. The Red force was fixed to represent three generic enemy vessel cases; slower (10 kt), similar speed (15 kt) and faster (25 kt). The following parameters and parameter ranges were identified as of being of interest:

Factor (units)	Min	Max	Step
Blue Manoeuvre Speed (kt)	15	25	5
Determination of intent range (m)	1,000	10,00	1,000
NLW range (m)	50	300	50
NLW hit probability	0.2	1	0.2
NLW effect duration (s)	4	12	2
Blue Manoeuvre Speed (kt)	15	25	5

Table 3: Data farming parameter ranges

This resulted in 4,500 distinct simulation models which were each run 30 times using a cloud computing cluster provided by NTU.

MEASURES OF EFFECTIVENESS

The principle measure of effectiveness used was whether Blue reached their goal. A success for Blue was if it could out-manoeuvre Red or keep Red at bay with its non-lethal weapon until Red ran out of motivation.

METHODLOGY

According to the parameters specified in Table 2, a gridded data farming experiment was conducted using OldMcData. The latter is a software application which was specifically designed (by NPS) to perform data farming experiments. Moreover, OldMcData has recently been adapted (by NTU and DSO) to support cloud computing infrastructures. This cloud computing approach enables one to utilise OldMcData on any cloud computing compliant clusters (e.g., the Amazon Elastic Compute Cloud). This novel approach facilitates the use of high performance computing facilities and addresses the user's computing scalability/budget issues.

The scenarios were executed on a Singaporean (NTU) cloud computer cluster based on the Apache Hadoop framework. The results were then analysed using the JMP statistical discovery software developed by SAS Institute Inc.

JMP's partition platform enables users to systematically analyse large data sets to discover relationships. JMP uses visualisation to create a successive tree of partitions according to a relationship between the X and Y variables. It finds a set of cuts or groupings of X values that best predict a Y value by exhaustively searching all possible cuts or groupings, recursively forming a tree of decision rules until the desired fit is reached.

RESULTS

Case 1: Red slower than Blue

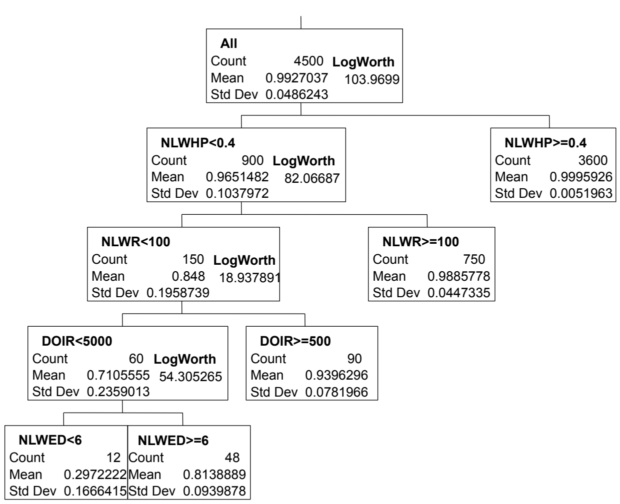


Figure 1: JMP regression tree for Case 1

The regression tree analysis shows that the critical parameters (in order) were the non-lethal weapon hit probability (NLWHP), non-lethal weapon range (NLWR), determination of intent range (DOIR) and non-lethal weapon effect duration (NLWED).

A non-lethal weapon effect duration of more than 6 seconds became critical in the event that Blue had a non-lethal weapon with a hit probability of less than 40%, a weapon range of less than 100 m and a determination of intent range less than 5 km.

Case 2: Red equal to or faster than Blue

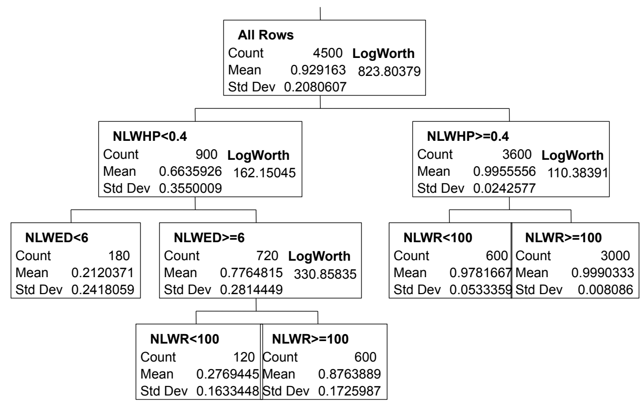


Figure 2: JMP regression tree for Case 2.1, Red speed equal to Blue.

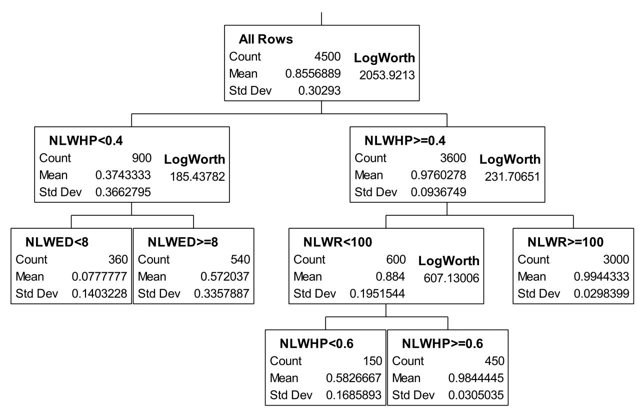


Figure 3: JMP regression tree for Case 2.2, Red speed faster than Blue.

This time the regression tree analysis shows that the critical parameters (in order and for both cases 2.1 and 2.2) were the non-lethal weapon hit probability (NLWHP), non-lethal weapon effect duration (NLWED), and non-lethal weapon range (NLWR).

The key difference between this and the previous case (where Red is slower than Blue) is that increasing the determination of intent range (DOIR) did not influence Blue's chance of survival.

Finally, we note that Blue's chance of survival diminishes as the speed of Red increases.

DISCUSSION

The high inertia of the Blue vessel meant it often failed to reach its maximum escape speed after recognising an attack was imminent. A better tactic for Blue may therefore be to transit at maximum speed the entire time.

This, along with other parameters, could have been explored in more detail, however time constraints dictated that this be done at a later date. Questions that could be explored further include:

- **Red motivation**: What if Red was willing to continue with their attack for longer than 15 minutes?
- **Red keep-out range**: What if Red was able to get Blue to surrender at longer stand-off ranges?
- **Initial Blue speed**: What if the Blue vessel starts off at faster speeds?
- **Blue NLW effect**: What if the non-lethal weapon was more or less effective at slowing Red down?
- **Blue inertia**: What if the Blue vessel was more manoeuvrable?

There was some discussion during the week about the MANA V inertia parameter needing to be quantifiable in future versions of MANA to be more useful (i.e. that is, have physical units associated with it).

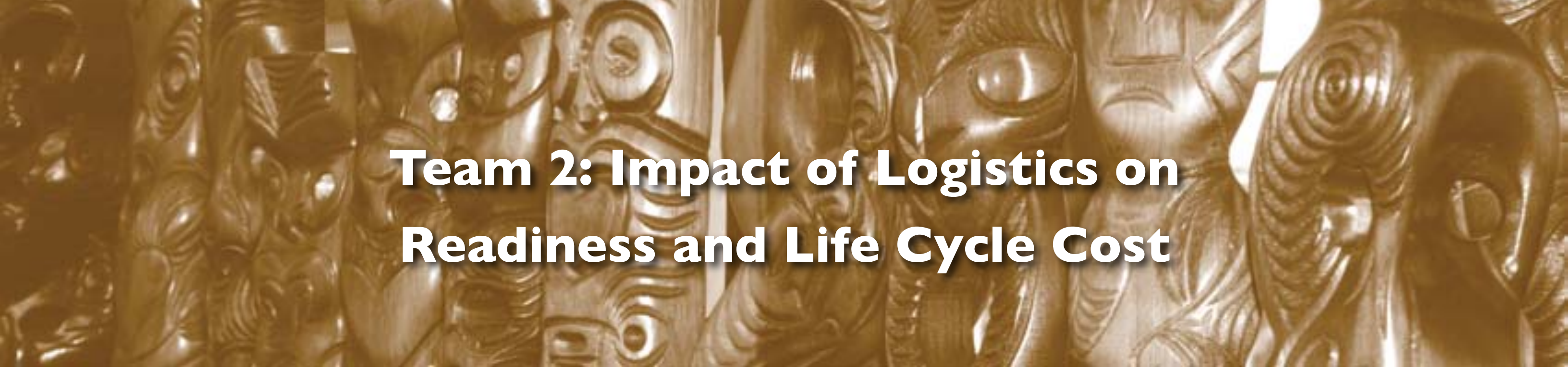
CONCLUSION

A long determination of intent range (in this particular scenario, greater than 5 km) was only found to be useful in the event that Blue had sufficient speed and acceleration to out-manoeuvre Red.

The following general insights into the non-lethal defence of a generic large vessel were obtained:

- Use non-lethal weapons that have a high hit probability (in this particular scenario, greater than 40% per shot).
- Use non-lethal weapons that have sufficient range (in this particular scenario, 100 m).
- Use a non-lethal weapon that has sufficient effect duration (in this particular scenario, greater than 6 seconds).

Overall, the results provide some interesting insights which will inform future analysis.



Team 2: Impact of Logistics on Readiness and Life Cycle Cost

TEAM 2 MEMBERS

Keebom Kang
Mary McDonald
Meredith Thompson
Naval Postgraduate School, US

Shawn Phillips, Captain
U.S. Marine Corps

Peng Kiang Low
Defence Science & Technology Agency, Singapore

Dawoomi Kim
Defence Technology Agency, New Zealand

INTRODUCTION

The development of modeling and simulation (M&S) to support Total Life Cycle Management (TLCM) is complex and requires a close examination of the factors involved and results obtained. The insights derived from M&S can contribute significantly to operational readiness and the cost of readiness. Consequently, research is needed to identify important data and key TLCM factors.

In this paper we develop two models that can be used to identify critical factors that impact on military readiness and the total life cycle cost. The first one is a simulation model using the Arena® simulation language (Kelton et al., 2010). This model estimates the operational availability of a major weapon system given input parameters under a certain scenario. The second one is an Excel® spreadsheet based life cycle cost model that computes the total life cycle cost using the same input parameters for the simulation model. Although the life cycle cost model can be embedded into the simulation model, we intentionally separated it utilizing the Excel spreadsheet, so that it can be used by financial managers to conduct what-if analyses virtually on any personal computer without obtaining a license or learning the simulation package.

We have selected the Light Armored Vehicle with a 25 mm Gun System (LAV-25) for our analysis. Our scenario includes 76 LAV-25s to be deployed with the Marine Expeditionary Forces. We are interested in readiness analysis for the LAV-25 fleet and in estimating the total life cycle cost over a period of time (typically 20 years). In this paper we only consider operations and maintenance cost without including infrastructure, personnel or overhead costs. The models discussed in this paper can be applicable to any major weapon systems by providing new data sets.

MODELS

Simulation Model for Operational Availability

There are 76 LAV-25s to be deployed with the Marine Expeditionary Forces. We are interested in readiness analysis for the LAV-25 fleet and in estimating a total life cycle cost over a period of time (typically 20 years). In this paper we only consider operations and maintenance cost without including infrastructure, personnel or overhead costs.

An LAV consists of 1,570 different parts. The data we collect from MCDSS indicate that the Marine Corps has spent a total of $29,372,715 to purchase all the parts to maintain LAVs in 2007 – 2009. After sorting the data in the order of the Extended Price (the total amount of money spent on each part; it is calculated by multiplying the part count by the unit cost), we note that more than $18 million or approximately two-thirds of the total cost were spent on the five parts listed in Table 1. These parts are expensive and fail more often than others. Also these parts are critical, i.e., if any one of these fails, the LAV cannot operate. We focus on these five parts for our analysis.

In Table 2, the failure rate, the SL quantity, the criticality code and the level of repair are shown for each part. According to MCDSS 4.3.1.1 Software User Manual, the failure rate is the number of failures per million days, which is incorrect. After consulting with the experts in the area, the time units were corrected to one calendar day instead of one million days. Even using one day as a time unit, the failure rates seem to be too low (i.e., or the parts are too reliable). In our analysis we did not use these failure rates given in Table 2. Instead we picked values from a wide range to conduct the design of experiments analysis using NOLH. More details on the design of experiments will be presented in the following sections. The criticality code (Crit Code in Column 5) 5 indicates all these parts are critical; any one of these parts fails the LAV cannot operate. The SL quantity (Column 4) identifies the total number of part(s) that are required on a selected weapon system. The SL quantity of the driving differential (Part #3) is 4, which is interpreted as 4 serially connected driving differentials in an LAV. We are assuming that if any one of them fails, the LAV will not be operational. If the reliability of an individual driving differential is 0.90, the subsystem that consists of these 4 serially connected driving differentials is only 0.656 (or 0.904). Thus among these 5 parts, the driving differential (Part #3) is potentially more critical to operational availability of the LAV than other parts. The SL quantity of the laser sensor unit (Part 1) is 2, which means that two of these parts are serially connected. This part can potentially make the critical part list.

The far right-hand-side column of Table 2 shows the level of repair for each component. Control display units (Part #2) and driving differentials (Part #3) are repaired at the intermediate-level (I-level) maintenance facility while laser sensor units (Part #1), and diesel engines (Parts #4 and #5) are maintained at the depot level (D-level). D-level as compared to I-level takes much longer turnaround time.

	Part Name	Part Count	Unit Price	Extended Price *
1	SENSOR UNIT,LASER	120	$89,794	$10,775,318
2	CONTROL DISPLAY UNIT	103	$27,683	$3,072,854
3	DIFFERENTIAL,DRIVING	101	$22,475	$2,269,978
4	ENGINE,DIESEL I	31	$41,757	$1,544,995
5	ENGINE,DIESEL I	36	$26,890	$1,075,609

Table 1: The Five Major Components for LAV (in terms of the Extended Price) [Source: MCDSS 4.3.1.1, PartUsage_EO947 (2007-2009)]

	Part Name	Failure Rate	SL Qty	Crit Code	Level of Repair
1	SENSOR UNIT,LASER	0.000211638	2	5	D-level
2	CONTROL DISPLAY UNIT	0.000363312	1	5	I-level
3	DIFFERENTIAL,DRIVING	8.90643E-05	4	5	I-level
4	ENGINE,DIESEL I	0.000109346	1	5	D-level
5	ENGINE,DIESEL II	0.000126983	1	5	D-level

Table 2: Failure Rate, SL Quantity, Criticality Code and Level of Repair for Each Part [Source: Same as TABLE 1]

We consider the five critical components as shown in Table 1 for this study. When any of these parts fails, the faulty part is removed from the LAV, an RFI (ready-for-issue) spare is installed, and the faulty part is sent to the repair facility (I-level or D-level). After the repair is complete, the repaired part becomes an RFI spare and is sent to the spare pool. When a critical part fails, and an RFI spare is not available, the LAV will be non-operational, or not mission capable, until an RFI part is available.

The ranges of the failure rates (λi) of the individual parts are provided in Table 3, along with the ranges of the number of spare parts, repair turnaround times (I-TAT and D-TAT), and the operational tempo. The failure rate (λi) is expressed in terms of the number of failures per operating hour. The minimum λi value of 0.0001 failures per operating hour is equivalent to an MTBF of 25 years (= 1 / 0.0001 / 400) assuming an operational tempo of 400 hours per year. The maximum value of 0.005 failures per operating hour is equivalent to an MTBF of 0.5 years (= 1 / 0.005 / 400) at the same operational tempo. The spare levels are set at 1 through 10. The I-level and D-level turnaround times are set at 5 – 15 days and 30 – 60 days, respectively. Since the average operational tempo during the peace time is approximately 350 hours per year, and during the war time, is approximately 650 hours per year, we set the range of the operational tempo from 300 to 700 hours per year. We assume that the time between failures for each component follows an exponential distribution. Our simulation model estimates the average operational availability. Our goal is to better understand how changes in reliabilities (in terms of the failure rate, λi), number of spares for each part, repair turnaround times, and operational tempo impact operational availability and the life cycle cost.

Input Parameter	Range
lambda i (λ_i, i = 1, 2, …, 5)	0.0001 – 0.005 failures per operating hr
spare i (i = 1, 2, …, 5)	1 - 10
I-TAT (I-level turnaround time)	5 – 15 days
D-TAT (D-level turnaround time)	30 – 60 days
Op Tempo	300 – 700 hrs/yr

Table 3. Ranges of Input Parameters

Several designs are possible, but we use an NOLH with 257 runs (Cioppa and Lucas 2006). This design is capable of handling up to 29 factors without increasing the number of scenarios. It can be easily constructed by entering the low and high values in Table 3 into a spreadsheet (Sanchez 2006). (We

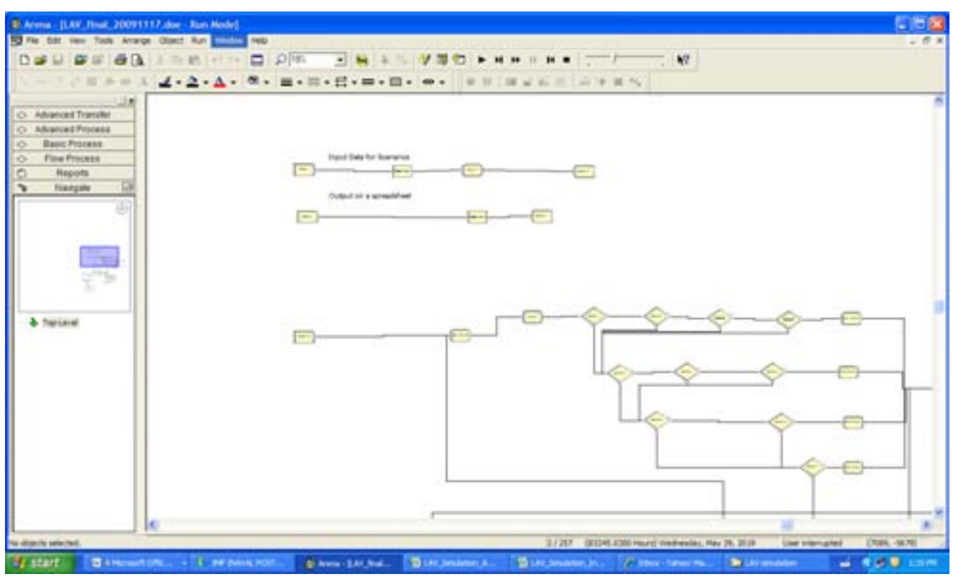


Figure 1. A Sample Screen Shot of the Arena Simulation Model

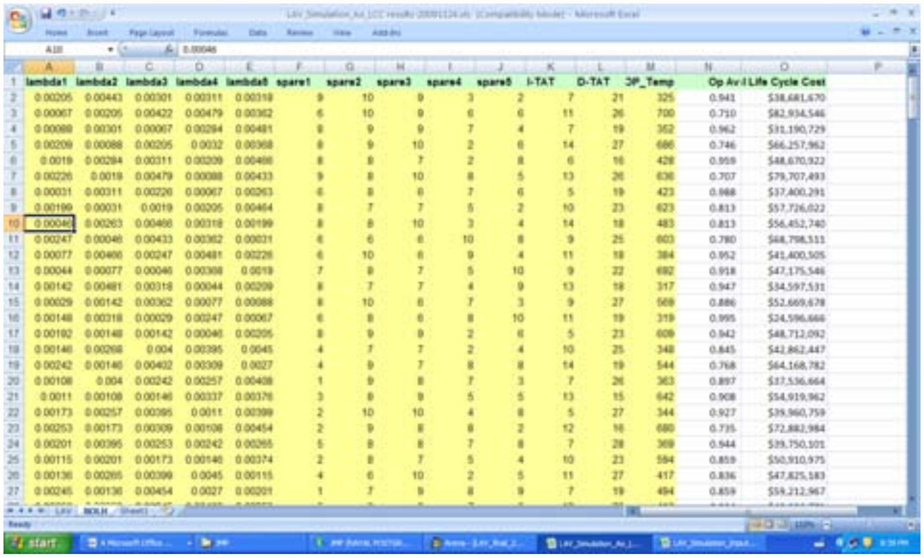


Figure 2. A Screen Shot of NOLH Input Parameters (Columns A through M) and Output Results (Ao in Column N and the life cycle cost in Column O)

remark that that 13 input factors could be examined using a NOLH with as few as 33 scenarios if the time required for 257 runs was prohibitively long.) Because our model runs quickly, we opt for a larger design to allow a more detailed investigation of our model's behavior. The input parameters for the first ten scenarios are shown in Table 4. In all, there are 13 different simulation input parameters used as factors for our designed experiment. In addition, there is a stochastic element that occurs due to the pseudo-random numbers generated for stochastic failure times, and repair turnaround times (I-level or D-level).

For each scenario, the simulation model reads a row of data from the spreadsheet excerpted in Table 4. The failure rates (λi, i = 1, 2, …,5) of the five parts are first read, followed by the number of spares for each part, the I-level turnaround time (I_TAT), the D-level turnaround time (D_TAT) and the operational tempo (Op_Temp). The λi values are expressed in terms of the number of failures per operating hour. The time between part failures is assumed to follow an exponential distribution for all 5 components we study. The repair turnaround times are assumed to follow symmetric triangular distributions with lower and upper bounds of 0.5(mean) and 1.5 (mean), respectively.

The simulation model was developed using the Arena simulation language. A sample screen shot of the Arena Simulation model is given in Figure 1. A brief description of the model logic is as follows:

1.Read input data for each scenario.

2.Generate 76 LAVs.

3.Generate part failures. (Five failure times are generated. Whichever is the smallest value is the next failure time of the LAV.)

4.The faulty part is removed from the LAV

5.Update FMC count (fully mission capable LAVs): FMC = FMC – 1

6.If a spare part is available, install a spare and reset the FMC count: FMC = FMC +1. Otherwise wait in the queue until a spare is available from the spare pool until after Step 8.

7.The faulty part is sent to the repair facility.

8. After the delay as specified in the input, the repaired part joins the spare pool.
9. Repeat Steps #1 to #8 until the end of the simulation time.
10. At the end each of scenario, the simulation automatically calculates the operational availability: Ao = average FMC / total number of LAV. FMC is a time-persistent variable (see e.g., Kelton et al. 2010) and the average value of FMC must be "time-averaged." Arena automatically computes the value.

We run a total of 257 scenarios, each of which is simulated over a period of 1,000,000 hours and the first

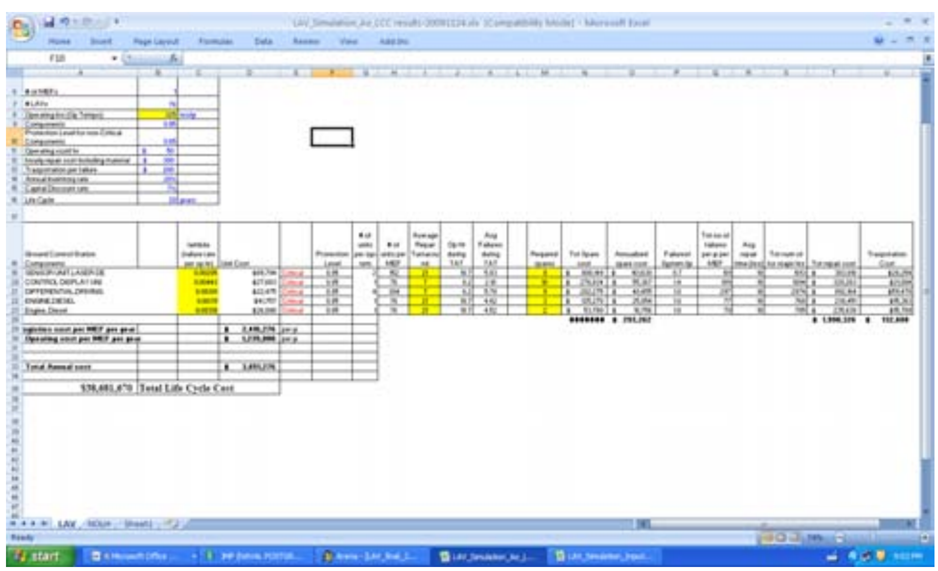


Figure 3. A Screen Shot of the LCC Computation Model

100,000 hours of observations were eliminated to remove initial bias. 900,000 hours of simulation is equivalent to approximately 7 replications of 20-year simulation. The average Ao (operational availability) from each scenario is automatically written onto the same input EXCEL spreadsheet worksheet (see Figure 2, Column N).

Spreadsheet Model for Total Life Cycle Cost

The second model is the life cycle cost spreadsheet model that reads the same input scenarios (Table 2) and computes the total life cycle cost over the next 20 years. The total life cycle in this model only includes operations and maintenance costs (spare, repair, transportation, and operations costs). See the sample screen shot in Figure 3.

Our scenario includes 76 LAVs. The operational tempo (Op Tempo or operating hours) changes during the design of experiments analysis. We assume that the hourly operating cost to be $50, the hourly repair cost, $300, and the transportation cost, $200 per failure. The annual capital discount rate of 7%, and the annual inventory rate of 20% are used. The life cycle is assumed to be 20 years for the LAV. Any of these values can be modified by the user.

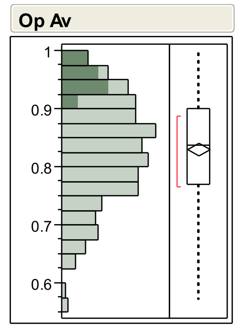
We have developed a visual basic macro program to automate the life cycle cost computation for each of 257 scenarios. Once the macro is executed, it reads the input parameters in each scenario (row by row) in Figure 2 then writes them on the highlighted cells on the spreadsheet life cycle cost model. Once the spreadsheet model updates the total life cycle cost, the result (cell A35) is written onto the same input worksheet (Column O). This procedure is repeated 257 times. When the macro execution is complete, the life cycle cost results are on Column O, and the operational availability results from the simulation are in Column N, along with input parameters in Columns A through M. [see Figure 2.] Then this worksheet is imported into the JMP® (SAS 2008) for further analysis.

RESULTS

We begin assessing the output by looking at histograms of the simulation responses. This can be a way of "accidentally" performing verification and validation of a simulation model. Our results indicate that the average operational availability differs widely across the different scenarios, ranging from 0.573 to 0.995. The average Ao across the 257 scenarios is 0.830 with a standard deviation of 0.090. It appears that at least one of the input factors does, indeed, have a substantial influence on the system's performance. [See Figure 4.]

The average total life cycle cost across the 257 scenarios is $52.1 million with a standard deviation of $14.6 million. The total life cycle cost also differs widely across the different scenarios, ranging from $19.9 million to $99.9 million. It also indicates that at least one of the input factors does have a substantial influence on the total life cycle cost. [See Figure 5.] The darkened areas of the Ao and cost distributions correspond to the subset of scenarios with lowest cost and highest Ao points, which will be explained in Figure 10 at the end of this section.

Op Av



Quantiles

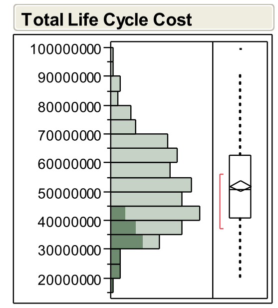
100.0%	maximum	0.99520
99.5%		0.99348
97.5%		0.98177
90.0%		0.94315
75.0%	quartile	0.89972
50.0%	median	0.83795
25.0%	quartile	0.76866
10.0%		0.70506
2.5%		0.64539
0.5%		0.57274
0.0%	minimum	0.57256

Moments

Mean	0.8297314
Std Dev	0.0897837
Std Err Mean	0.0056006
Upper 95% Mean	0.8407604
Lower 95% Mean	0.8187023
N	257

Figure 4. Distribution of Operational Availability (Op Av)

Total Life Cycle Cost



Quantiles

100.0%	maximum	99541745
99.5%		97157458
97.5%		85228536
90.0%		71038691
75.0%	quartile	62690658
50.0%	median	50571562
25.0%	quartile	40754869
10.0%		34565056
2.5%		26239206
0.5%		21212985
0.0%	minimum	19915925

Moments

Mean	52125320
Std Dev	14641807
Std Err Mean	913330.84
Upper 95% Mean	53923918
Lower 95% Mean	50326721
N	257

Figure 5. Distribution of Total Life Cycle Cost

After confirming that the results appear reasonable, we turn to our main goals—identifying those factors and components that have the greatest impact on performance. A useful non-parametric tool is a regression tree, as shown in Figure 6. These graphics are effective for understanding and communicating the results of thousands of runs over many factors. Regression trees are more human-readable and can be easier to describe than multiple regression models because they reveal the structure in the data in a simple way. Initially,

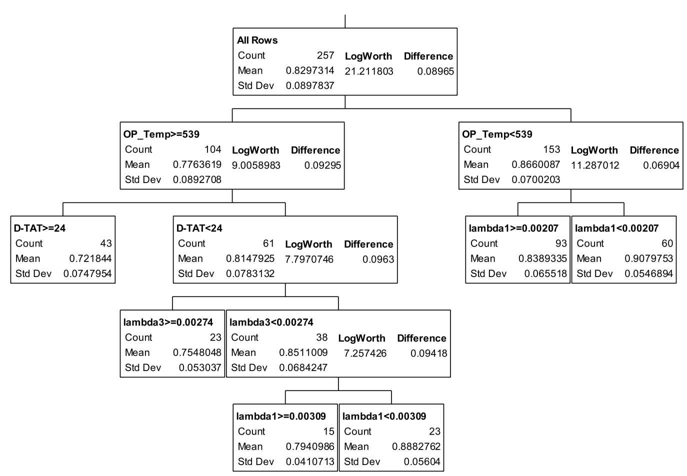


Figure 6. Regression Tree for the Average Ao

the data are grouped in a single cluster. All potential input factors are examined to identify how best to split them to yield two leaves so that the variability in the response within each leaf decreases and the variability in the response between the leaves increases.

Figure 6 shows the regression tree for predicting the average Ao from the 257 simulation scenarios. The dominant factor is clearly the operational tempo (Op_Temp). For example, the first split at the top indicates that the average Ao is 0.776 across the 104 scenarios that have an operational tempo of 539 hours or more per year. In contrast, the average Ao is 0.866 (11.6% higher) among the 104 scenarios that had an operational tempo less than 539 hours per year. As the operational tempo increases (i.e., more operating hours), the more failures occur and obviously the operational availability goes down. In the second split when the operational tempo is greater than 539, the depot turnaround time (D-TAT) becomes a critical factor. Then the failure rates (lambda 1 and lambda 3) of the laser sensor unit (Parts #1) and the driving differentials (Part #3) become critical.

Regression trees are non-parametric approaches for fitting a statistical model to the simulation output. They can be good at identifying subsets of the output that behave much differently than the rest. Regression metamodels can also be valuable. They may confirm the regression tree results concerning which factor or factors have the greatest influence on the results, or they may allow more succinct descriptions of the simulation model's performance if it can be well-described by simple polynomial metamodels.

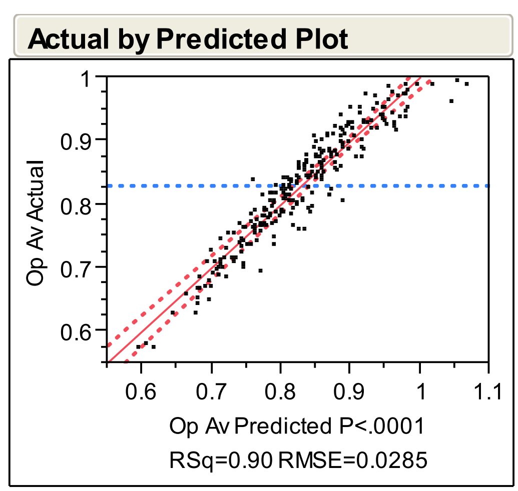
Accordingly, we fit regression metamodels of the Ao as a function of main effects, and two-way interactions of the 13 input factors. After noticing the impact of spare parts on readiness is not significant, we fit regression metamodels with 8 input factors and two-way interactions (excluding spares from the main effects.) The sorted parameter estimates from the JMP analysis are shown in Figure 7 with R2 of 0.98 (not shown here), indicating that the regression metamodel does an excellent job of explaining the variability in the simulation output. The results also confirm the observations made on the previously discussed regression tree.

Because it can be difficult to look at a regression equation and get a good sense of how the factors and interactions affect the response, interaction plots are often useful. This interaction plot consists of several small subplots that indicate how the predicted performance (Ao) varies as a function of pairs of input factors. After examining the interaction plots (not shown here), we do not see any major impact of interaction effects on the operational availability other than the interaction between lambda3 and I-TAT, implying that the impact of high failure rate of the driving differential (Part #3) is mitigated by small I-TAT. (Note that Part #3 is maintained at the I-level as previously discussed.) However, this interaction effect is not significant compared to the main effects. Hence, we develop a simple linear model with 13 main factors without interaction effects. The results are shown in Figure 8 with R2 of 0.90, which might also be used to make inferences. The parameter estimates are shown in Table 5. The large |t_ratio|s for the OP_Temp, lambda1, lambda3 and D-TAT show them to be the major factors, and agree with our regression tree results. Note that the numbers of spares are not as important as other factors. This means that raising the spare levels from their lowest levels to the highest levels in Table 1 does not lead to appreciable improvement in the average operational availability.

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
OP_Temp	-0.000418	1.875e-5	-22.29	<.0001*
lambda1	-27.47609	1.530956	-17.95	<.0001*
lambda3	-24.93189	1.531073	-16.28	<.0001*
D-TAT	-0.006014	0.000499	-12.05	<.0001*
I-TAT	-0.00682	0.000744	-9.17	<.0001*
lambda4	-8.939229	1.531038	-5.84	<.0001*
lambda5	-8.751261	1.530919	-5.72	<.0001*
(lambda3-0.00255)*(I-TAT-10.0039)	-2.799943	0.504818	-5.55	<.0001*
(lambda3-0.00255)*(OP_Temp-500.031)	-0.041816	0.015023	-2.78	0.0058*
(D-TAT-22.5019)*(OP_Temp-500.031)	-1.352e-5	4.917e-6	-2.75	0.0065*
(lambda1-0.00255)*(OP_Temp-500.031)	-0.035695	0.013523	-2.64	0.0089*
(lambda1-0.00255)*(D-TAT-22.5019)	-0.989734	0.416004	-2.38	0.0182*
(lambda2-0.00255)*(D-TAT-22.5019)	0.892101	0.392613	2.27	0.0240*
(lambda3-0.00255)*(D-TAT-22.5019)	0.788151	0.375328	2.10	0.0369*
(lambda1-0.00255)*(lambda3-0.00255)	2408.4318	1150.557	2.09	0.0375*
(lambda5-0.00255)*(D-TAT-22.5019)	-0.702307	0.353283	-1.99	0.0481*
(lambda4-0.00255)*(I-TAT-10.0039)	1.0563666	0.588199	1.80	0.0739
(lambda4-0.00255)*(OP_Temp-500.031)	-0.020614	0.012939	-1.59	0.1126
(I-TAT-10.0039)*(OP_Temp-500.031)	-1.123e-5	7.055e-6	-1.59	0.1129
lambda2	-2.121909	1.530933	-1.39	0.1671
(lambda3-0.00255)*(lambda4-0.00255)	1474.5518	1219.623	1.21	0.2280
(lambda2-0.00255)*(OP_Temp-500.031)	-0.013942	0.014265	-0.98	0.3295
(lambda4-0.00255)*(D-TAT-22.5019)	-0.382095	0.423	-0.90	0.3674
(I-TAT-10.0039)*(D-TAT-22.5019)	0.0001377	0.000178	0.77	0.4408
(lambda5-0.00255)*(I-TAT-10.0039)	0.4185263	0.615759	0.68	0.4974
(lambda2-0.00255)*(lambda5-0.00255)	928.85083	1382.665	0.67	0.5024
(lambda3-0.00255)*(lambda5-0.00255)	709.11581	1086.993	0.65	0.5148
(lambda2-0.00255)*(I-TAT-10.0039)	-0.331117	0.814144	-0.41	0.6846
(lambda4-0.00255)*(lambda5-0.00255)	-443.6789	1255.473	-0.35	0.7241
(lambda1-0.00255)*(lambda2-0.00255)	-440.9017	1303.218	-0.34	0.7354
(lambda5-0.00255)*(OP_Temp-500.031)	-0.004375	0.014702	-0.30	0.7663
(lambda1-0.00255)*(lambda4-0.00255)	268.16666	987.1597	0.27	0.7861
(lambda1-0.00255)*(I-TAT-10.0039)	-0.139088	0.721916	-0.19	0.8474
(lambda1-0.00255)*(lambda5-0.00255)	317.03533	1713.308	0.19	0.8534
(lambda2-0.00255)*(lambda3-0.00255)	-47.32484	1146.104	-0.04	0.9671
(lambda2-0.00255)*(lambda4-0.00255)	44.091172	1193.08	0.04	0.9706

Figure 7. Sorted Parameter Estimates from JMP Analysis



Figures 8. Actual by Predicted Plot of Operational Availability (Linear Regression)

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	1.3106713	0.017487	74.95	<.0001*
lambda1	-27.52971	1.252032	-21.99	<.0001*
lambda2	-1.864504	1.252123	-1.49	0.1378
lambda3	-24.78179	1.253018	-19.78	<.0001*
lambda4	-8.967459	1.252148	-7.16	<.0001*
lambda5	-8.489004	1.252103	-6.78	<.0001*
spare1	0.0069739	0.000673	10.36	<.0001*
spare2	0.0007581	0.000673	1.13	0.2611
spare3	0.0056174	0.000673	8.35	<.0001*
spare4	0.0039842	0.000673	5.92	<.0001*
spare5	0.0034526	0.000673	5.13	<.0001*
I-TAT	-0.006878	0.000608	-11.31	<.0001*
D-TAT	-0.006007	0.000408	-14.72	<.0001*
OP_Temp	-0.000417	1.534e-5	-27.21	<.0001*

Table 5. Parameter Estimates for Ao

We use the same methodology to analyze the results from the total life cycle cost model. We first observe the regression tree as shown in Figure 9. Once again the dominant factor is the operational tempo (OP_Temp). For example, the first split at the top indicates that the average life cycle cost is $63.5 million across the 114 scenarios that had an operational tempo greater than or equal to 523 hours per year. In contrast, the average life cycle cost is $43 million among the 143 scenarios that had an operational tempo less than 523 hours per year. As the operational tempo increases, the operating cost obviously increases, and this factor has more impact on the life cycle cost than any others. In the second split, when the operational tempo is greater than 523, the reliability of the driving differential (Part #3) becomes a critical factor.

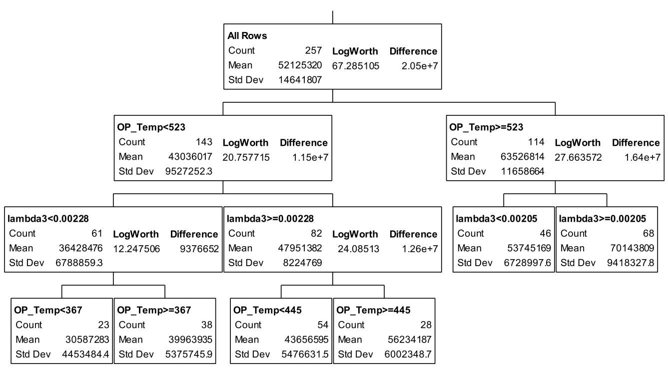


Figure 9. Regression Tree for the Total Life Cycle Cost

We then develop regression metamodels with two-way interaction effects. After observing that not many interactions are significant, we build a simple linear regression model, and the results are shown in Table 6. The R-square value is 0.98 showing that this linear regression model does an excellent job of explaining the variability in the life cycle cost model. The results shown in Table 6 confirm that the operational tempo is the most critical factor followed by lambda3 and lambda1 (the failure rates of Part #3 and #1). However, the repair turnaround times (I-TAT and D-TAT) do not affect much on the total life cycle cost. As the repair turnaround times increase, the operational availability deteriorates, yet the operations and maintenance cost will not be affected. However, there is an implied waste of money since the repair pipeline inventory will go up as the repair turnaround time increases.

The number of spares for Part #1 (the most expensive one among the five parts selected for analysis) seems to be more critical than other spares, yet the number of spares in general does not have much impact on the life cycle cost.

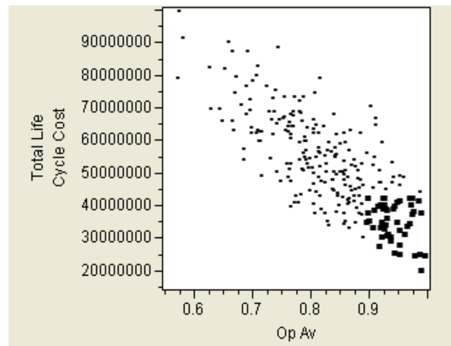


Figure 10. Scatterplot of Cost vs Ao

Lastly, we compare Ao and total life cycle cost together in a scatterplot, given in Figure 10. The darkened points correspond to those scenarios with lowest cost and highest Ao. There is an obvious negative correlation between the two measures, with higher operational availability generally leading to lower overall total life cycle cost. In examining the input factor distributions for these selected points (not shown here), we find that the factor with the most explanatory power for achieving both lower cost and higher Ao together is operational tempo. The next most driving influence for both measures considered together is the failure rate of Part # 3. These darkened points are also shown as the darkened areas in the distribution graphs of Figures 4 and 5.

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-29561567	1254676	-23.56	<.0001*
lambda1	2.5765e+9	89829872	28.68	<.0001*
lambda2	1.2881e+9	89836403	14.34	<.0001*
lambda3	5.1531e+9	89900658	57.32	<.0001*
lambda4	1.2884e+9	89838195	14.34	<.0001*
lambda5	1.2885e+9	89835006	14.34	<.0001*
spare1	190172.1	48278.49	3.94	0.0001*
spare2	58604.502	48287.37	1.21	0.2261
spare3	47574.053	48258.82	0.99	0.3252
spare4	88359.169	48268.53	1.83	0.0684
spare5	57063.63	48257.42	1.18	0.2382
I-TAT	-596.3134	43644.69	-0.01	0.9891
D-TAT	-48.45175	29278.29	-0.00	0.9987
OP_Temp	99386.948	1100.438	90.32	<.0001*

Summary of Fit

RSquare	0.981489
RSquare Adj	0.980498
Root Mean Square Error	2044709
Mean of Response	52125320
Observations (or Sum Wgts)	257

Table 6. Parameter Estimates and a Summary of Fit for the Total Life Cycle Cost Model

REMARKS

The most critical factor is the operational tempo for both Ao and the total life cycle cost. The more you operate, the more it fails and the more it costs. Then the reliability plays an important role. Those parts with high failure rates are more critical (e.g., Part #1 and Part #3). D-level turnaround times as well as I-level turnaround times are also critical factors for Ao. Those with longer repair turnaround times [e.g., Part #1 fails less often than Part #3, yet its repair turnaround time is much longer (D-level maintenance) than that of Part #3 (I-level maintenance)] tend to be more critical for Ao. The repair turnaround times do not have much direct impact on the total life cycle cost, since, in the life cycle spreadsheet model, the spare level for each component was predetermined for each scenario. However, in reality if the repair turnaround time gets longer, managers tend to purchase more spares to improve readiness, yet those newly acquired spares will only spike the operational availability for a short term. Eventually those newly acquired spares will fail and get stuck in the repair pipeline without improving Ao in the long run if the repair turnaround time is not reduced.

In conclusion, both warfighters and financial managers should understand the importance of logistics and impact of reliability, cycle time and operational tempo on readiness and life cycle cost.

REFERENCES

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

[2] Kelton, W. D., R. P. Sadowski, and N. B. Swets. 2010. Simulation with Arena, Fifth Edition, McGrow-Hill.

[3] 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.

[4] SAS. 2008. JMP® User's Guide, Version 8: SAS Institute Inc. Cary, NC.

Team 3: Analyzing Selected Questions in a Refugee Camp Scenario Using PAX3D

TEAM 3 MEMBERS

Seichter, Stephan LtCol
Bundeswehr Transformation Center, Germany

Lampe, Thorsten
Maly, Andreas
EADS System Design Centre, Germany

TIAH Yao Ming
Defence Science Organisation, Singapore

Prof. Sanchez, Susan
Hernandez, Andrea
SEED Center for Data Farming, USA

Grieger, Dion
Defence Science Technology Organisation, Australia

Silwood, Narelle
Defence Technology Agency, New Zealand

INTRODUCTION

The Bundeswehr Transformation Center is examining how M&S can effectively support CD&E projects related to Peace Support Operations (PSO). Human Factors and Human Behavior analyses have proven to be highly relevant in this context. One study specifically examines possibilities to model scenarios in a PSO with PAX in which the military is tasked to assist in building and operating refugee camps, and especially to ensure order and security. PAX is planned to be used to support decision makers in assessing and evaluating ROE (Rules of Engagements) applying the basic use of force guidelines for the soldier in PSO missions.

Questions

The model of evolving aggressiveness within groups which is modeled in PAX is quite complex (a subsumption can be found in [1]).

Therefore, one major object of investigation for the team was which input parameters that determine the group's behavioral characteristics lead to escalating and particularly dangerous, i.e. violent situations. This question will be stated more precisely when looking at the specific scenario and studies later on. However, since there are a lot of factors actually characterizing the behavior of human agents with regard to their emotional states and evolvement of aggression, statistical means will be used to examine those factors and their interrelationship in our studies.

At IDFW18, PAX was calibrated in specific micro scenarios within a refugee camp scenario modeled in the 3-

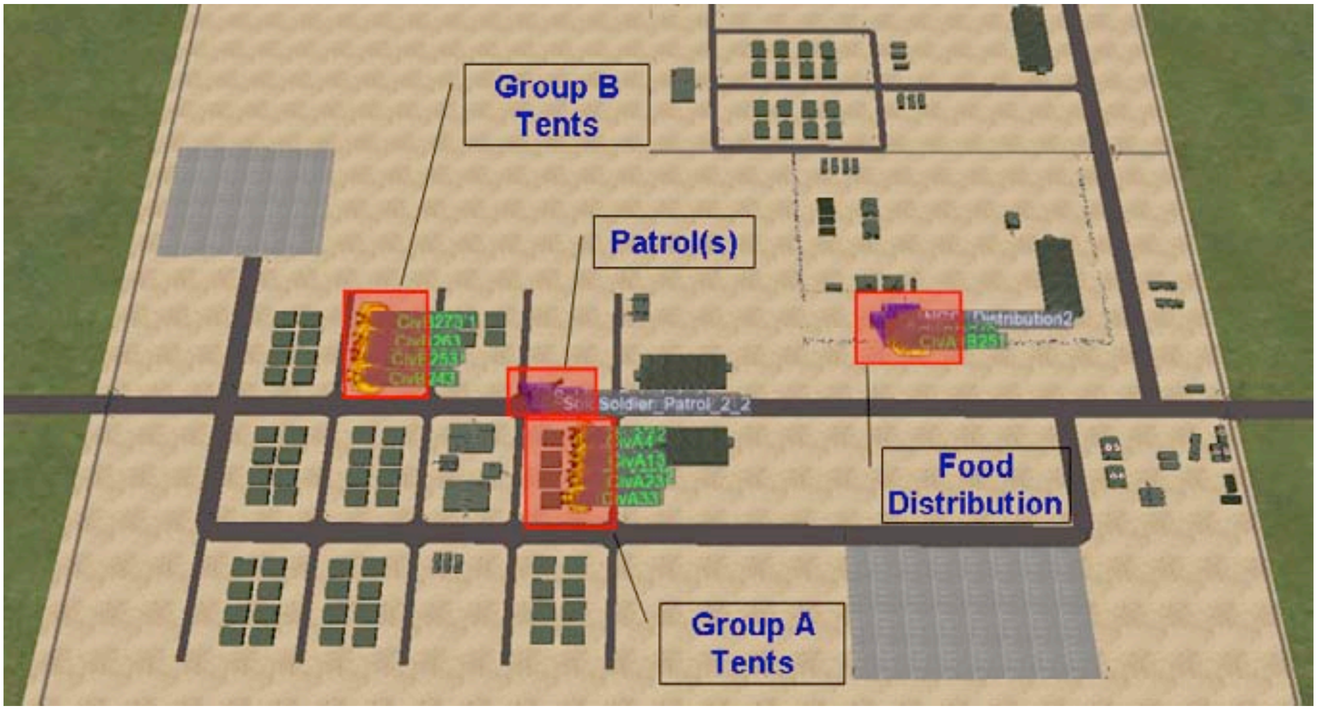


Figure 1: Refugee Camp Base Scenario

dimensional environment of PAX. Special focus was on calibrating the soldiers' behavioral model in detail, which had been adopted to represent a realistic self defense behavior of the soldier agents according to commonly applied use of force guidelines.

Having found that the soldiers' self defense behavior is carried out correctly and in a sufficiently realistic way in the micro scenarios with very few agents, the team's second objective during IDFW19 was to find out whether this soldier behavior also works properly in larger ("macro") scenarios.

To summarize, the driving questions for investigation for the week were:

- Which group characteristics lead to escalating and dangerous situations? Which input factors and combinations thereof are most influential?
- Does the new soldier behavior that was validated during IDFW18 in micro scenarios also work properly in larger scenarios?

Scenario

Figure 1 depicts the general refugee camp scenario that was used as the baseline for the week's studies in its implementation using the current PAX3D version. The scenario is set up to be at around noon time at the time when food distribution at the central distribution point has just started and representatives of each tent are in the area marked as Food Distribution to collect the respective tent's packages.

Meanwhile, the rest of the refugees are waiting in their tents for the tent representatives to bring back and distribute the food. While this is their main motivation, there is also a certain tension between the two rivaling groups marked as Group A and Group B with 40 refugees each. Besides the "normal" civilians, each group encloses 2 disturbers, initially aggressive and equipped with weapons. A soldier patrol is set up in the tent area to prevent or dissolve trouble amongst the refugees.

STUDIES & ANALYSIS

To address our questions and in particular to examine the most influential factor combinations for certain group characteristics and scenario developments, we started off with a Data Farming experiment in which we varied the parameters characterizing the dynamics of PAX with regard to the agents' emotional states.

Experiment 1 Setup

While the complete design in terms of the varied range of each factor is shown in Table 1, it is important to briefly introduce the meaning of the parameters, at least to an extent which will support the following analysis:

A civilian's anger factor (denoted by PC_ANGER in the following) determines the dynamics of the anger of an agent. A low anger factor results in an angry agent staying angry for a long time, thus characterizing – in combination with other model inputs – a rather hot-tempered civilian.

	Min	Max
A Normal PC_ANGER	0.01	0.2
A Disturbers PC_ANGER	0.01	0.2
B Normal PC_ANGER	0.01	0.2
B Disturbers PC_ANGER	0.01	0.2
A Normal PC_FEAR	0.01	0.7
A Disturbers PC_FEAR	0.01	0.7
B Normal PC_FEAR	0.01	0.7
B Disturbers PC_FEAR	0.01	0.7
A Normal PC_RFA	0.01	1.0
A Disturbers PC_RFA	0.01	1.0
B Normal PC_RFA	0.01	1.0
B Disturbers PC_RFA	0.01	1.0
A Normal PC_AROUSAL	1.05	2.0
A Disturbers PC_AROUSAL	1.05	2.0
B Normal PC_AROUSAL	1.05	2.0
B Disturbers PC_AROUSAL	1.05	2.0
Soldiers' Sensor Range	10.0	50.0
Threshold Anger Very High	71.0	95.0
Threshold Anger High	50.0	70.0
Threshold RfA Very High	71.0	95.0
Threshold RfA High	50.0	70.0

Table 1: Basis of NOLH design of experiment

Likewise, a low fear factor (PC_FEAR) of a civilian characterizes a rather anxious personality.

A similar correlation exists for the readiness for aggression (RFA) and the arousal of the agents, respectively.

The soldiers' sensor range determines the range in which the soldiers react to interactions, representing a very basic notion of a "sensor".

Finally we varied some internal thresholds for classifying anger and readiness for aggression as high and very high, respectively, during a simulation run. They are displayed in the table but will not be examined further.

Experiment 1 Analysis

When examining this first NOLH experiment, one much unexpected observation was that in some of the runs a very high number of civilians (up to a maximum of 19) were actually killed (see Figure 2).

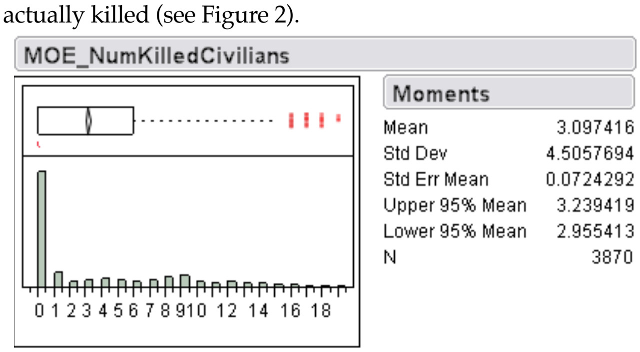


Figure 2: Distribution of number of killed civilians in NOLH design experiment

Although this was a surprise at first, this outcome actually helped the team tremendously in answering our first question. We noticed that the very diverse spectrum of scenario outcomes was mainly due to the wide ranges in which we had varied our model factors. As we had done so intentionally in order to identify which factor ranges and their combinations caused different types of group behavior, we now took a closer look at the combinations that caused this high level of aggression.

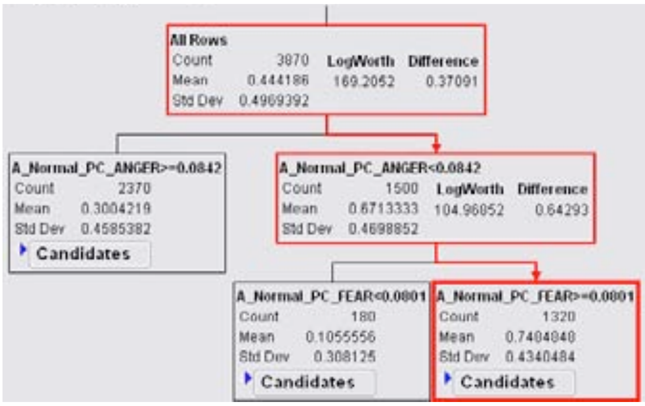


Figure 3: Regression tree analysis of MOE "At least one civilian killed"

A regression tree analysis showed that the combination of the PC_ANGER and the PC_FEAR parameters is a key factor for the violence observed in the scenario. For example, Figure 3 shows that a rather low anger factor (meaning the agents stay rather angry) combined with a high fear factor (meaning that the agents are rather "fearless") is a key influence for at least one civilian being killed.

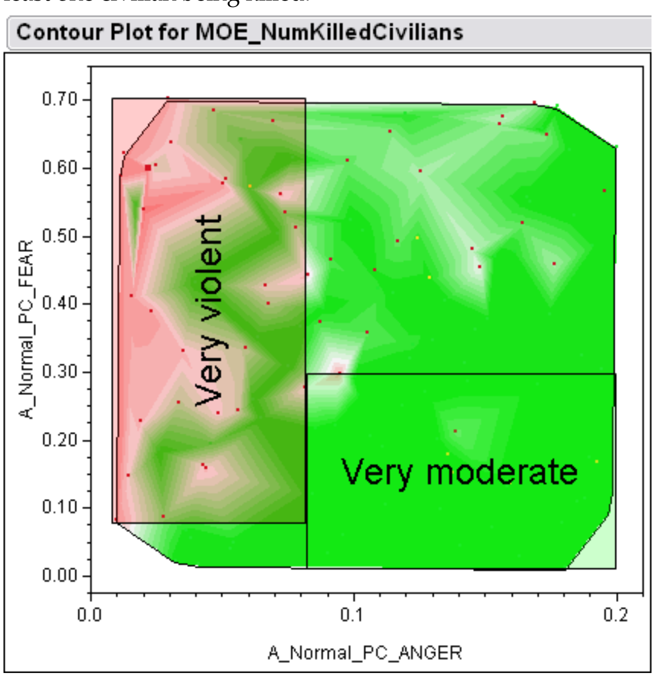


Figure 4: Allocation of moderate ⊠ violent civilian personality characteristics (green = 0 kills to red = 19 kills)

In other words, we can conclude that a combination of an anger factor below 0.08 and a fear factor above 0.08 leads to the agents behaving very aggressively in this scenario. On the other hand, we found that with an anger factor above 0.08 and a fear factor below 0.3 very few incidents with killed civilians happen. Figure 4 displays these cross-effects graphically, although these results will have to be confirmed by follow-up experiments.

One can further interpret Figure 4 concluding that the remaining parts of the parameter space characterize medium or normal aggression levels of the respective civilians once more thorough investigation of the two determining factors has been done.

As to the team's second question we identified a remaining flaw in the implementation which in our scenario setup caused the soldiers not to react strongly enough as long as other civilians were beaten, but not attacked with any weapons. Especially in the very violent scenarios this had the effect of aggressors continuously beating members of the other group, unimpressed by the soldiers' continuous attempts to resolve the situation.1

Experiment 2 Setup

In a follow-up experiment we re-ran the same experimental design with a slightly changed base case scenario in which in addition to the disturbers 5 randomly selected "normal" members of each civilian group were equipped with weapons. The team hoped that this would lead to the soldiers recognizing the life-threatening situation, solving the arising conflicts.

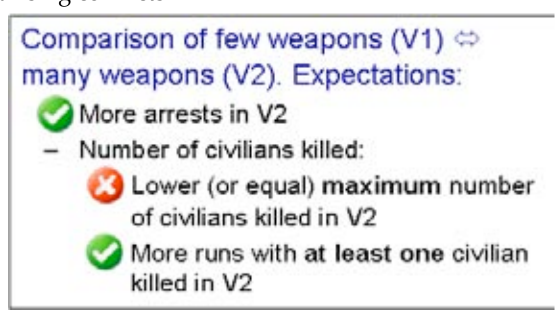


Figure 5: Team expectations in second experiment

The team's expectations as to the outcome of this scenario are summarized in Figure 5, together with a first overview of the results of the analysis. The figure shows that the team expected fewer civilians to be killed in the second experiment, which is an intuitive expectation when considering the aforementioned reason for setting up the second experiment with more weapons.

Experiment 2 Analysis

Two of our expectations were met in the experiment analysis: The civilians' carrying more weapons did actually lead to more of them being arrested and there were more runs with at least one civilian killed.

However, the fact that the soldiers now consequently performed their self-defense behavior and arrested disturbers

[1] The soldiers' behavior is not introduced here, but a description can be found in [2].

did not in turn lead to fewer civilians killed in general. We noticed that the effect of more extreme attacks happening outweighed the soldiers' stepping in more thoroughly.

The comparison of the contour plots of the experiments with (Figure 6) and without weapons clearly confirms this observation. Also, in the "medium-range" area (upper right) we can see more severe and violent simulation runs than in experiment 1 – an effect of the higher number of weapons being used by the civilians.

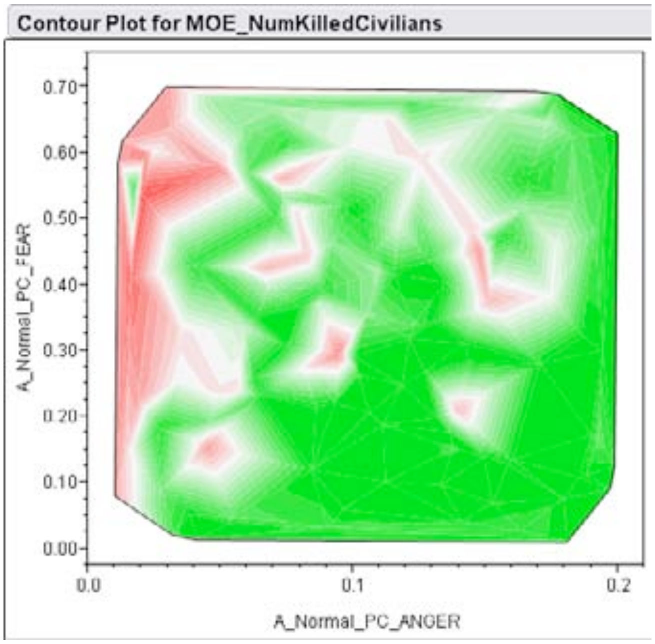


Figure 6: Influence of PC_ANGER and PC_FEAR in 2nd experiment (green = 0 kills, red = 20 kills)

Thus in essence the rough parameter ranges of PC_ANGER x PC_FEAR that we had identified in the first experiment were confirmed to hold in the vignette with weapons as well.

We conclude that we have made a step forward in understanding the complex dynamics of the model and finding parameter ranges that allow us to model different group characteristics.

On the other hand, the soldiers' inability to control the situation in the violent scenarios shows the necessity to make the soldiers' behavior suitable also for larger scenarios with many incidents happening at the same time. This will require some coordination within and between different soldier groups and their respective leaders, not yet implemented in PAX, as well as a slightly enhanced vulnerability model allowing for the soldiers to better assess the harm that an aggressive action implies to the victim.

CONCLUSION

With regard to our goal of determining relevant factors for modeling different group characteristics in PAX3D we have found through our experiments that we are able to model a wide variety of different scenarios – from very peaceful scenarios to very violent ones.

Furthermore, we were able to identify distinct combinations of some of the PAX3D personality constants, among them especially the anger and fear personality constants, leading to extreme scenarios with regard to violence.

In follow-up experiments we will examine these parameter combinations in more depth with the goal of eventually being able to predefine specific agent templates with respect to the agent's or group behavioral characteristics – for example very violent and aggressive, moderate or peaceful civilians and groups. These "soft" attributes, while intuitive to the human reader (at first), are generally not easy to transfer into the hard technical variables of a simulation model. Therefore, we consider the work of IDFW19 a major step forward in this direction.

As a side effect of our studies during the week, it has once again become evident that in accordance with the question-based paradigm of the Data Farming concept a simulation model is and has to be built to answer specific questions. In our case, we conclude that when in the course of the scenario the situation turns into a combat-like situation, a model other than PAX3D might be better suited to represent this new situation, focusing on the combat aspects in addition to the social aspects of the scenario.

The Bundeswehr Transformation Center is currently investigating how far functionalities of combat-oriented models can be introduced into PAX to enhance the model in this direction, and to identify the cutting point where a scenario really requires a different model. Furthermore, in an operational synthesis approach as investigated by our team during IDFW16, PAX and more combat-oriented models could eventually complement each other in these types of scenarios.

In terms of further calibration of the model and scenario we found that the soldiers' behavior, while working well in small-scale ("micro") scenarios, needs to be further adapted for macro scenarios where many incidents happen at the same time. A basic notion of coordination between the different soldier patrols and their leaders, in our case, is seen as the foremost necessity for modeling this type of scenarios in order to retain a more realistic and sensible soldier behavior.

The results of IDFW 18 could be proven practically feasible for model calibration. The distinction of the model factors into three categories (see [2]) makes the calibration process more targeted: Internal factors are the "deepest" in the model and are to be calibrated to fixed values, ideally never touched again. Advanced factors are calibrated to fixed values or ranges, but the values depend on the scenario and thus requiring recalibration when the scenario changes. The variable factors finally represent the parameters available to the OR analyst and should be calibrated to reasonable ranges to achieve model feasibility depending on the scenario being analyzed. This categorization is considered essential for future calibration work.

To summarize, we have shown that to some extent we can classify the civilian group characteristics into rough templates and have identified parameter ranges for each template – moderate to highly aggressive – which will have to be narrowed down and divided further in the future. Beyond that, the soldiers in PAX will need additional abilities, such as coordination of squads or more realistic sensing and a situational awareness, which have both shown to be essential in larger scenarios.

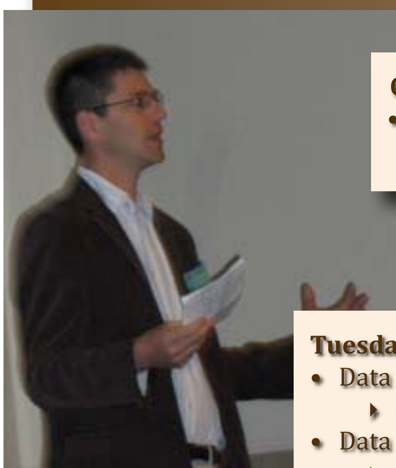
Finally, the interdisciplinary, international and collaborative atmosphere during IDFW19 again guaranteed great work with valuable results! Special thanks to all team members for bringing in their expertise, work and time, and fun during the week – with one word: their MANA!

REFERENCES

[1] Schwarz, G., and Mosler, H.-J. Investigating Escalation Processes in Peace Support Operations: An Agent-based Model about Collective Aggression. In Proceedings of the 3rd Annual Conference of the European Social Simulation Association (ESSA 2005), ESSA, Koblenz, Germany, 2005.

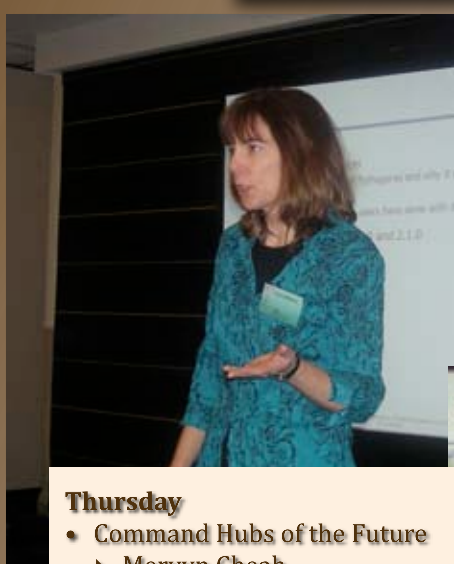
[2] Lampe, T., and Seichter, S. PAX3D Refugee Camp Scenario – Calibration of the Adapted PAX Model. Team 05 report in Scythe Proceedings and Bulletin of the International Data Farming Community (Scythe 6), IDFW18, Monterey, CA, USA, 2009.

Plenary Sessions



Opening - Keynote Speaker
- David Galligan
  - New Zealand Defence Force

Tuesday
- Data Farming for New Members
  - Gary Horne
- Data Farming Tools
  - Steve Upton and Ted Meyer
- MANA Version 5
  - Gregory McIntosh
- Pythagoras Update
  - Donna Middleton



Wednesday
- Design of Experiments
  - Susan Sanchez
- A Model for Representing Strategies for Causing Changes in Allegiances within a Population
  - Michael Lauren
- Automated Red Teaming Framework
  - Xu Yong Liang
- Evolvable Simulations and Cloud Computing
  - James Decraene

Thursday
- Command Hubs of the Future
  - Mervyn Cheah
- Discrete-Event Modeling
  - Paul Sanchez

Team 4: Evaluation of electro-optical sensor systems in network centric operations using ABSEM 0.3

TEAM 4 MEMBERS

Karsten Haymann - Lead
Dr. Daniel Nitsch
Federal Office of Defense Technology and Procurement, Germany

Dr. Gary Horne
Naval Postgraduate School, USA

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

INTRODUCTION

The development of the agent-based sensor and effector model ABSEM was started in 2008 by EADS on behalf of the German Federal Office of Defence Technology and Procurement. Since then it has been continuously enhanced and at IDFW19 version 0.3 was released. 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 new model version is characterized by more sophisticated effector modeling and extended possibilities regarding the setup of the agents' behavior. Furthermore it contains a first radar model implementation.

During IDFW19 a camp protection scenario was used to perform extensive data farming experiments and thus improve system understanding.

Objectives

Using a camp protection scenario, the team's objective is to investigate the effect of different sensor systems (electro-optical and radar) and effector types (direct / indirect fire, weapons with point and area effect) within a dynamic environment. Particularly the new ABSEM version 0.3 features were to be reviewed. This included on the one hand verifying the integrated model approaches and on the other hand exploring the model's possibilities when it comes to simulating more complex NCO scenarios. The new features are:

- Radar systems for airborne reconnaissance
- Larger terrain cell (400*400km), urban environment
- IEDs, indirect fire, weapons with area effect
- Extended state dependent
  agents' behavior state machine
- Trigger events

In Data Farming experiments the team's main intention is to examine the effect of the given sensor and effector systems under varying conditions, such as different weather-dependent atmospheric conditions, time of day, varying number and type of blue and red units,…).

Overall, the team has the following goals:

- Review and face validate ABSEM version 0.3
- Validate the implemented radar model
- Setup a complex scenario
- Conduct data farming experiments analyzing the effect of parameters such as different sensor systems, number of deployed blue forces, reaction times,…
- Find out further model and data farming requirements

Simple scenario for testing the radar model's plausibility

A new radar model has been implemented for the new ABSEM version 0.3. At this workshop a first simple radar scenario was simulated to investigate and validate the correct modeling of monostatic pulsed radar systems in ABSEM.

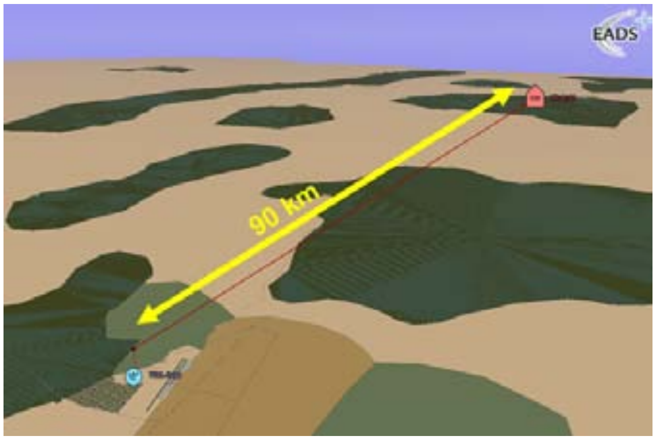


Figure 1: Simple radar test scenario

In the scenario shown in figure 1, an aircraft is approaching an air surveillance radar starting at a distance of around 90 kilometers.

The detection distance was used as the MOE to measure the radar performance for different radar types, weather conditions and target sizes and types.

The implemented radar model considers the following parameters:

First, a wide range of technical radar parameters may be defined and used to simulate the radar performance in the radar model.

Second, the environment is taken into account (e.g. line of sight or atmospheric losses due to rain).

And finally target parameters like the Radar Cross Section (RCS) and the Swerling Case, which defines the fluctuation of the RCS is taken into account in the radar model.

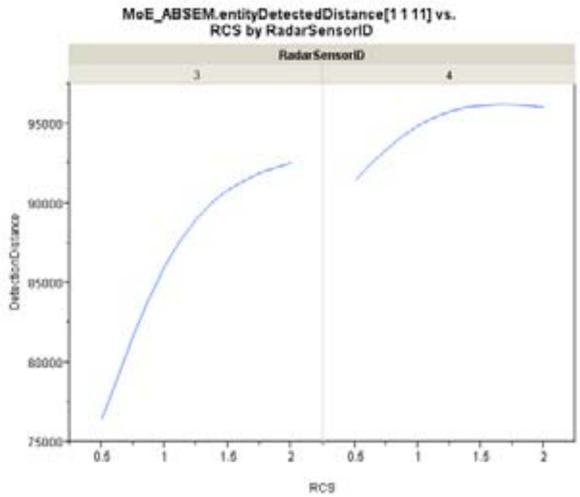


Figure 2: Detection distance for different RCS values for low (left) and high (right) frequency radars

Overall we could verify the radar model's correctness. The performed data experiments confirmed the expected behavior.

We could for instance observe that radars with higher frequencies are more robust against RCS-fluctuations (see figure 2) or technical radar losses (see figure 3). Lower frequency radars, however, perform better when it's very rainy. When looking at the regression tree we found out that besides the RCS value (which only depends on the target) the pulse compression and technical radar losses have the most influence on the radar performance.

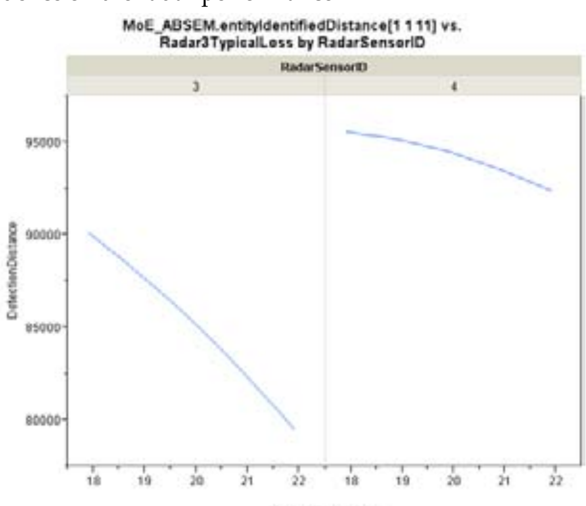


Figure 3: Detection distance for different typical radar losses for low (left) and high (right) frequency radar.

More complex NCO Scenario

The extended model functionalities regarding both the implemented effector model and the enhanced agent behavior were tested using a camp protection scenario.

Scenario Description

The military camp is located outside Mazar-e-Sharif and faces a constant threat by local insurgents. Once in a while the camp is attacked by mortar grenades mostly fired from the inner city. For this reason patrols through Mazar-e-Sharif are trying to detect and defeat any hostile firing positions. In addition UAVs are deployed for aerial surveillance to reconnoiter the attackers. All the reconnaissance information is passed on to the camp's headquarters.

In case a hostile mortar firing position was identified, all soldiers within the camp will be warned which causes them to retreat into secure shelters. Additionally the patrols will get the order to fight the insurgents. This implies that the blue patrol will quit its predetermined patrol route and follow or move towards the red insurgents. Those, however, have planned to perform a hit and run ambush by using an IED and attacking the patrol with rifles and rocket propelled grenades as soon as one of the patrol's vehicles was damaged by the IED and therefore the whole patrol was halted.

Figure 4: Overall scenario from the bird's eye view

Data Farming Experiments

We were executing a series of data farming experiments, looking at the following parameters:

- number of deployed UAVs for airborne reconnaissance: {0;1;2}
- deployment of a second vehicle patrol: yes/no
- time of the day: noon / midnight
- weather: foggy / clear
- type of sensor system used by blue forces: normal viewing during day and night / long wave infrared device

As MoEs we were mainly looking at the damage state of the blue forces, differing between the losses within and those outside the camp.

All of our experiments were successfully executed on the 32-node German cluster owned by 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 more than 5000 simulation runs, successfully executed in several hours on the 32-node German cluster owned by the German Procurement Office.

Data Farming Results

We found out that in this scenario setup the deployment of UAVs is essential for the camp protection, whereas the convoy patrols are not sufficient.

An interesting effect, however, can be seen in figure 5. Whereas the UAVs help a lot regarding the protection of the soldiers within the camp, it also leads to more losses in the convoy. The reason for that is that as soon as the red forces were identified by the UAV, the blue patrols will try to fight and follow the red forces. Unfortunately that's exactly what the red forces were trying to achieve and helps them ambushing the patrol. Thus, if the red forces were not detected, the patrol won't be involved in any fight.

Furthermore we could observe that actually the time at which the red mortar starts firing at the military camp is quite significant. However, this is a factor that cannot be influenced in reality. Therefore we were interested to find out which of the other parameters mostly affects the number of blue damages. To do so we generated a partition tree that showed that the number of deployed UAVs has the largest impact, followed by the type of sensor applied.

But it was also seen that apparently in an urban environment the deployment of highly developed sensor systems is not that important since the target needs to be very close anyway in order to be detected.

SUMMARY AND WAY AHEAD

The implemented radar model delivered very plausible results, thus showing us that we are on the right track.

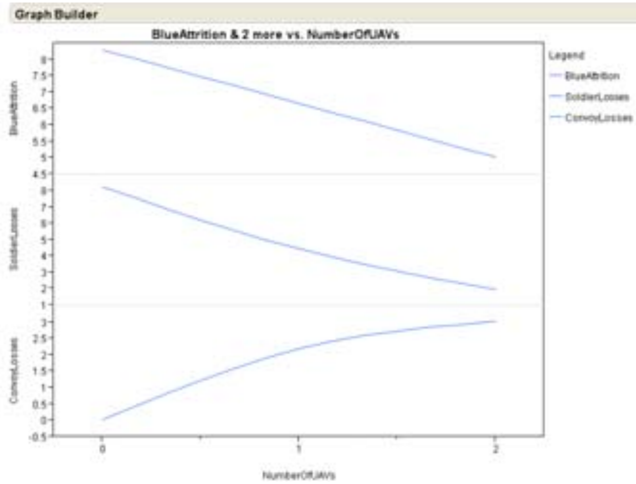


Figure 5: Influence of the UAV deployment on the number of blue losses

With ABSEM version 0.3 and the provided model features we are now able to set up more complex scenarios within a short amount of time.

Of course, the more complex the scenario and the longer the period of time we are looking at, the longer the execution times of the simulation runs. Therefore in future we will have to think about using more sophisticated experiment designs than just using the gridded design. In ABSEM, however, for the current set of analyzed scenarios we need to have the possibility to lockstep several parameters (for instance to model a convoy existing of several entities). Therefore we need to find a possibility for both using the NOLH design but still being able to lockstep. Additionally we derived the need for further user interfaces simplifying the whole scenario setup and the agents' behavior parameterization in particular.

Team 5: Force Protection with ITSim (II): Base Protection Against Ballistic Weapons

TEAM 5 MEMBERS

Pfeiffer, Volker M.
IT-AmtBw, Germany

Mausberg, Niklas
Mies, Christoph
Hörcher, Sascha
Wolters, Benjamin
Fraunhofer IAIS, Germany

Sanchez, Paul PhD
Naval Postgraduate School, US

Schubert, Johan PhD
Swedish Defence Research Agency, Sweden

Ruiz, Juan Dr.
Modelling and Simulation Coordination Office, NATO

INTRODUCTION

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 often hard to determine whether located humans are opponents (red) or just civilians (neutral). We use a base protection scenario and evaluate several active defense options against small teams firing improvised ballistic rockets at the camp. Based on the scenario introduced at the International Data Farming Workshop 18 (IDFW18) [1], the Force-Protection domain is enhanced to investigate further issues.

The investigated scenario analyzes exactly that aspect by using 3D terrain augmented with semantic information provided by the German Armed Forces. The data is not modeled but imported from an official data source. During this workshop, we wanted to answer two questions:

1. Does the consideration of semantic information (see below) result in a statistically significant change of the investigated Measure of Effectiveness (MoE)?
   Note that the consideration of semantic information will result in a more realistic environment model. But this more precise model raises costs in computation and modelling time. If the MoE is not affected by this additional effort, we can omit semantic information for this scenario.
2. Does an optimization of blue emplacements in order to increase the observed area (see below) result in a statistically significant improvement of blue's success w.r.t. the investigated MoE?

Of course, we expect the answer to this question to be true. It is interesting to investigate the importance of the optimization criterion w.r.t. the investigated MoE. If the criterion is not important, the MoE will not be affected. In future, we are interested in performing several optimizations according to several criteria in order to determine the most important ones.

Figure 1: Base in 3D terrain with semantic information

SCENARIO

Figure 1 depicts the investigated scenario. A blue base is located in 3D terrain with additional semantic information. Dark regions mark high terrain elevation whereas bright areas denote lower terrain. Thus, the blue base is located on a hill. The semantic information is attached directly to the terrain data. Basically, it is a classification of the terrain, including rivers, buildings, different types of wood, different types of streets and flat terrain. In figure 1, woods are visualized as green areas and plain terrain is depicted in yellow. Additionally, the dark roadmap and the blue river can be recognized. In the lower right part of figure 1, many buildings colored in red are visible. Two towers equipped with cameras are used to observe the surrounding area of the camp. They are visualized by tactical icons in the upper part of figure 1. During the course of the scenario, some Red will approach the base in order to attack it with ballistic weapons.

The key idea is that the opponents cannot be detected as Red until they start to prepare their attack. Thus, the whole approach time cannot be used to prevent the attack. After the configured preparation time, the opponents launch n missiles (with reload time in between) and flee afterwards.

The scenario's analysis is divided into two phases. The first one is a static classification and the second one is a simulation capturing the dynamics of the strategies.

Static Classification

Before the scenario is simulated dynamically, a static classification is performed. Two important measures are vital for the strategies: Ballistic threat and line-of-sight. Areas from which the base can be attacked by ballistic weapons are called ballistically threatening. The muzzle velocity of the weapon defines its maximal distance. The terrain defines if there is an angle that results in a flight trajectory such that the base can potentially be hit. Considering the semantic information, we defined that it is impossible to attack ballistically from rivers, buildings and woods. Thus, the ballistic threat strongly depends on the given terrain and semantic augmentation. The line-of-sight denotes which areas can be observed by the cameras in the base. These cells are called observable.

In order to perform the classification, the area around the base is gridded. Afterwards, every cell, i.e. grid element, is checked if it is ballistically threatening and observable. Note that the terrain itself is not gridded but based on precise vector data. According to that classification, three cases exist:

- *Green*: A cell is not ballistically threatening, i.e. the base cannot be attacked from that cell. The Blue don't have to worry about that cell. Therefore, the cell is colored green.
- *Yellow*: A cell is ballistically threatening and observable. Thus, the base can be attacked from that cell and there is a line-of-sight to the Blue. The attackers can be detected while they prepare their attack. The cell is colored yellow.
- *Red*: A cell is ballistically threatening and not observable. Thus, the base can be attacked from that cell and there is no line-of-sight to the base. The attackers cannot be identified while they prepare their attack. This is the worst case for the blue forces and the cell is colored red.

If the semantic information is taken into account in the grid classification, the ballistic reachability is restricted by not allowing shots from woods, rivers and buildings. This reduces the opponent's area of operation. This might have an effect on the optimization, because a smaller area needs to be monitored. Thus, fewer emplacements might be needed in order to establish a certain success for Blue. The result of this classification is depicted in figures 2 and 3. Figure 2 shows the classification with considering the semantic information, figure 3 without. Considering the semantics changes the number of green, i.e. not ballistically threatening, cells from 21.8% to 48.2%. Thus, roughly 26.4 percent of the cells are not longer threatening for Blue. This consideration of ballistic semantics results in a limitation for Red's area of operation. By comparing figures 1, 2 and 3, we can recognize that the woods (depicted green in figure 1) are ballistically threatening in figure 3 (classification without semantics) but are not in figure 2 (classification with semantics). The same holds for the river, which can be recognized in figure 2. Note that when considering a higher grid resolution, i.e. smaller grid cells, the whole calculation gets more accurate and the river can be recognized more clearly.

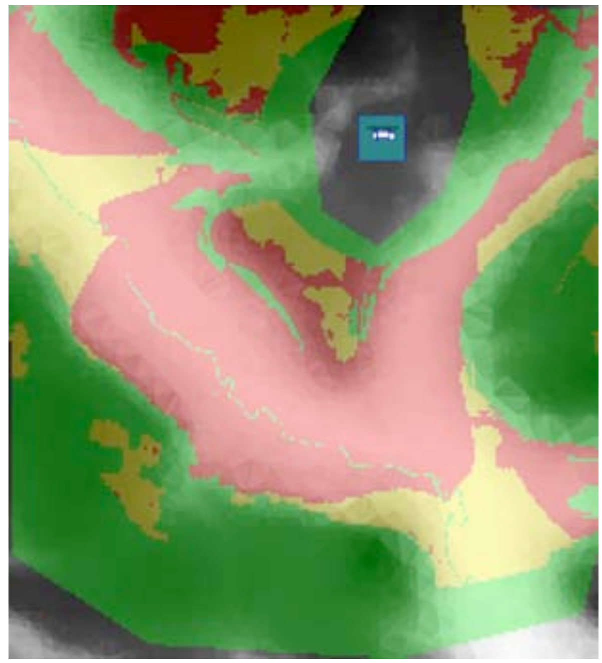


Figure 2: Result of classification considering semantics

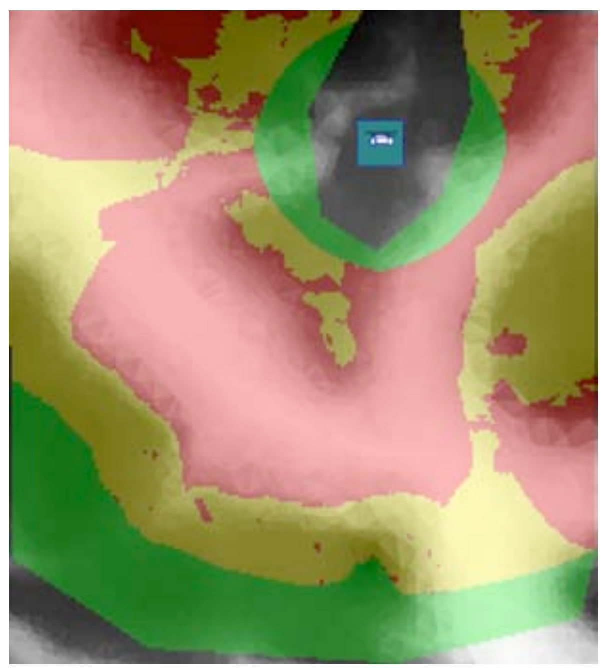


Figure 3: Classification without considering semantics

Simulation of the Strategies

In order to be able to compare the semantics impact on Blue's success, several strategies have been evaluated against a given red behavior. This kind of analysis may give interesting hints to support the defending of the base. The red strategy is fixed in all experiments. It consists of the following steps:

- *Generation*: The units are generated uniformly distributed outside the base. Their affiliation is neutral, i.e. they cannot be detected as hostile.
- *Approach*: A yellow or red cell (i.e. a ballistic attack is possible from that cell) is selected and moved to. The unit is still not detectable as hostile.
- *Preparation*: Two cases exist. If the attacker can detect any blue unit it gets discouraged and flees. Otherwise it starts to prepare its attack. From that point in time, it can be detected as hostile by Blue. As soon as the blue force is detected by the red unit, the Red aborts its preparation and flees. Note that the cameras' sight range is much higher than the one for regular ground troops including red attackers and blue defenders.
- *Attack*: The Red starts to fire a previously defined number of projectiles (intended shot number) at the base. From this point in time, the attacker is detected as hostile by the blue defenders if it has not already been. Between the shots, the attacker has to reload. Afterwards, it flees.

Currently, Blue has three different strategy options to prevent ballistic bombardment at the base:

1. *Pursue from Base* (PfB): A blue Quick Reaction Force (QRF) is located inside the base and pursues the red attackers as soon as they have been detected. The attacker can be observed by the cameras or they reveal themselves by shooting projectiles at the base.
2. *Camouflaged Emplacements* (CE): Camouflaged spotters are located outside the base. They can detect the Red but not vice versa. As soon as the red units are located, their position is reported to the base and the QRF starts the counterattack at the Red.
3. *Show of Forces* (SoF): Patrols move around the base. They can detect the Red and can also be detected by these. If any red force is located, the nearest patrol starts a counter attack. Note that there is no QRF in the base as in the other strategies. A red opponent detecting an approaching patrol is going to flee.

The semantic attributes influence the routing of the units. According to the different road types, a restriction to the top-speed along the roads is modeled. Thus, Blue's as well as Red's speed will be reduced on certain routes. One might expect less success for Blue, but this is not clear since both parties drive slower.

As a MoE we used the percentage of PreventedShots, defined as the ratio of prevented shots with respect to the number of intended shots. For example, if the red attacker intended to shoot two times and has been neutralized after one shot, PreventedShots is 0.5. Additionally, we measured the following MoEs, which are not further regarded in this report due to space limitations:

- *PreventedShots*: The number of prevented shots at the base. This happens if the attacker is neutralized or discouraged before the attack is started.
- *PreventedAllShots*: This binary MoE is true, when all intended shots have been prevented.
- *NeutralizedAttacker*: The number of neutralized attackers.

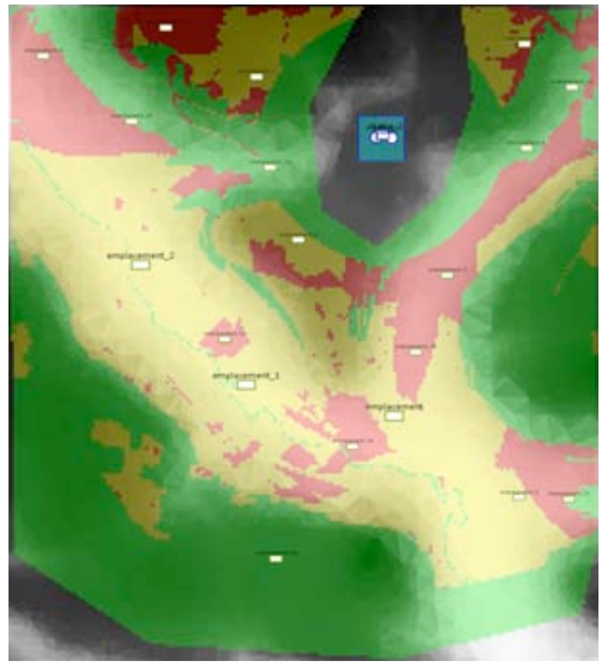


Figure 4: Classification with 3 active emplacements considering ballistic semantics

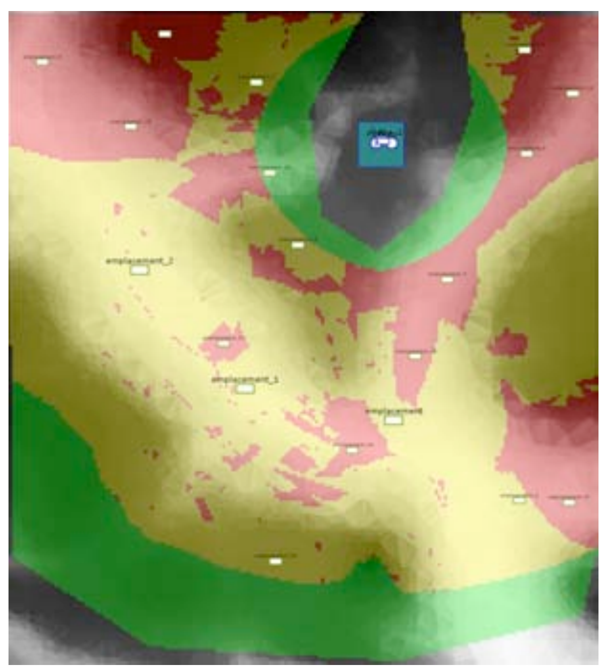


Figure 5: Classification with 3 active emplacements without considering semantics

In order to distribute the emplacements in the CE vignette, an optimization feature is used. The user has to define several possible positions and ITSim distributes up to n spotters over these possible positions. Figures 4 and 5 depict the possible emplacement position, the green icons. The visibility of the 3 active emplacements (the larger icons) is currently considered in the classification.

Figure 4 shows the optimal distribution of 3 emplacements considering the ballistic semantics and figure 5 shows the same distribution on a grid without considering ballistic semantics. The green cells remain unchanged as can be seen by a comparison of figure 4 and 5 with 2 and 3, respectively. Considering the semantics, the number of green cells is 18,144 of 37,638, i.e. 48.21 percent. When omitting the semantics, we only have 8,209 corresponding to a percentage of 21.81. The optimization maximizes the number of yellow cells with constant green ones. Thus, the optimization criterion is the number of yellow cells divided by the sum of red and yellow cells.

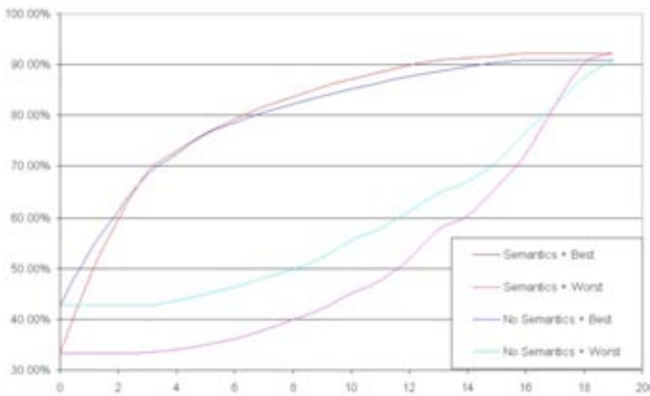


Figure 6: Percentage of yellow cells

Figure 6 shows this measure for the best and worst distribution of 0 up to 19 emplacements with and without considering semantics. The typical logarithmic shaped curve indicates that saturation is reached. The best distributions are much better than the worst ones if the number of possible decisions is sufficiently high. Note that the 0- and 19-distribution are exactly the same in the worst and best optimization since no decision can be made. Running the optimization on the grid considering ballistic semantics gives better results if at least six emplacements are distributed than the optimization results on the non-semantic one. The advantage of considering the semantics is because of the restricted operational area for Red. Note that the non-semantic grid is superior without any spotters being distributed. The reason for this is that most cells getting green are also visible by the watch towers. In eight of nineteen distributions, the result is different.

In SoF, the QRF is not waiting in the base but patrolling in the valley as depicted in figure 7. The blue line is its current route. The red attacker in the upper part of figure 7 is currently approaching its improvised fire position. In this case, a QRF positioned in the base would probably perform better. Note that the camera towers inside the base always support the detection of the Red. As mentioned above, the red units can only be detected after they have started preparing their attack. The QRF has limited time to reach the attackers before they can fire their rockets.

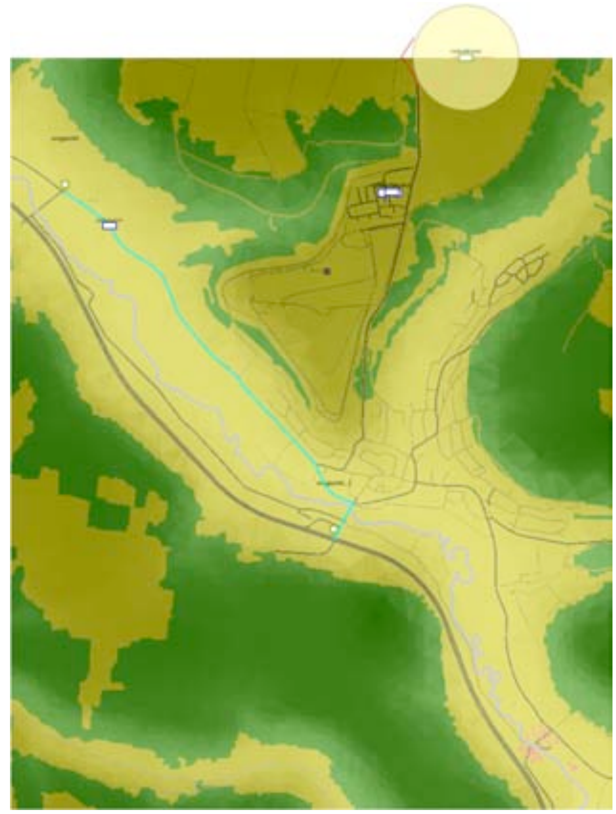


Figure 7: Running simulation of SoF

Although we defined three different strategy options for the blue forces, we are not interested in comparing these. As already stated in the introduction, we want to evaluate the influence of the additional semantic information as well as the optimization feature. Basically, PfB is a sub-strategy of CE (with no emplacements) as well as SoF with only one patrol staying in the base. In the following, we want to use all strategies to evaluate the semantics' influence on the MoE. With the CE strategies, we want to determine the impact of the optimization feature.

Factor	Min	Max	Unit
Speed *QRF*	20	75	km/h
Speed Red	20	75	km/h
Height Red	1.5	3.0	m
Setup Time Red	2	10	min
Reload Time Red	1	3	min
Mean Detection Time Camera	1	3	min
Mean Detection Time Emplacements	1	3	min
Intended Shots Red	1	3	

Table 1: NOLH design

RESULTS AND ANALYSIS

We used a Nearly Orthogonal Latin Hypercube (NOLH) [2] design with 65 design points for our experiments and crossed it with the two kinds of semantics. Thus, all experiments have been run on the grid depicted in figures 2 and 3, with and without considering the semantics for the ballistic reachability, respectively. Additionally, we ran all scenarios by using the semantic routing or not. By using it, the routing algorithm is aware of the top-speed restriction on the roads. The NOLH design (see table 1) consists of the following parameters: Speed of blue QRF, mean time of detection of the cameras on the watch towers in the blue base, the mean time of detection of the blue emplacements, the speed of the red attackers, the height of Red, Red's setup time (the time needed to build up its ballistic weapon), the reload time of Red and the number of intended shots at the base. For each design point 50 replications are conducted with different seeds. In total 13,000 simulation runs for each scenario are performed. Altogether, we modeled 8 scenarios, i.e. one PfB, one SoF and six CE scenarios. The latter scenarios were calculated with 1, 3 and 5 spotters, distributed according to best and worst optimization results, respectively. Thus, we performed 104,000 simulation runs.

As mentioned above, we wanted to examine the impact of considering the semantics as well as performing the optimization on blue's success, i.e. the MoE PreventedShots. Additionally, we compare Blue's strategies very briefly.

Impact of Optimization

The optimization result depends on the ballistic specific interpretation of the semantics. If the information is considered, different distributions are calculated as already mentioned above and can be seen in figure 6. For simplicity, we only used the results of the grid considering the semantics and evaluated this distribution with both grids, i.e. the semantic and non-semantic one. We distributed one, three and five emplacements according to the best and worst optimization result. Tables 2 and 3 show the results. Note that 0 observers correspond to the PfB strategy serving as baseline. The worst distribution of observers made the MoE PreventedShots even worse.

# observer	Ballistic Semantics		No Ballistic Semantics	
0	mean	0.331	mean	0.205
	std-dev	0.435	std-dev	0.362
	std-err	0.005	std-err	0.004
1	mean	0.359	Mean	0.264
	std-dev	0.442	std-dev	0.405
	std-err	0.005	std-err	0.005
3	mean	0.470	mean	0.360
	std-dev	0.468	std-dev	0.441
	std-err	0.006	std-err	0.005
5	mean	0.561	mean	0.446
	std-dev	0.463	std-dev	0.462
	std-err	0.006	std-err	0.006

Table 2: "Best" optimization of CE

In most cases, the best distribution outperformed the worst one significantly. Only the non-ballistic grid with one emplacement is an outlier. The reason therefore is that the optimization has been performed considering ballistic semantics but the simulation has not. Thus, the optimization criterion does not match the simulated reality. This artifact is a hint that the model of the optimization is not accurate enough and is a great indication that accurate and correct models must be used during optimization and simulation in order to get robust results. Thus, we claim that our optimization according to the visibility of the spotters improves Blue's statistically significant.

# observer	Ballistic Semantics		No Ballistic Semantics	
0	mean	0.331	mean	0.205
	std-dev	0.435	std-dev	0.362
	std-err	0.005	std-err	0.004
1	mean	0.316	mean	0.288
	std-dev	0.428	std-dev	0.416
	std-err	0.005	std-err	0.005
3	mean	0.235	mean	0.219
	std-dev	0.389	std-dev	0.375
	std-err	0.005	std-err	0.005
5	mean	0.203	mean	0.241
	std-dev	0.374	std-dev	0.390
	std-err	0.005	std-err	0.005

Table 3: "Worst" optimization of CE

Impact of Semantics

"Does considering the semantics have an impact on Blue's success?" - "It depends."

The impact of the semantic information is different according to the different effects. We modeled two effects: The impact on the ballistic reachability as well as the impact on the semantic routing referred to as ballistic semantics and road semantics, respectively.

Strategy	Ballistic Semantics		No Ballistic Semantics	
PfB	mean	0.331	mean	0.205
	std-dev	0.435	std-dev	0.362
	std-err	0.005	std-err	0.004
SoF	mean	0.532	mean	0.536
	std-dev	0.463	std-dev	0.471
	std-err	0.006	std-err	0.006

Table 4: PreventedShots of PfB and SoF

The impact of the ballistic semantics can be read off tables 2, 3 and 4 for the different strategies. Especially with CE, a significant effect can be seen, since in all cases, Blue's success raises when ballistic semantics has been considered during simulation. The main reason therefore is that the semantics has been considered during the optimization step. For the baseline PfB, the impact is also significant. The reason therefore is that many cells that are green, i.e. not ballistically threatening, only if the semantic is considered are located far away from the base (cf. figures 2 and 3). These cells cannot be reached by the QRF in time with high probability. Thus, Red can launch more attacks if the ballistic semantics is not considered. Using strategy SoF Blue performs not significantly different in comparison to the scenario without semantics. The reason is that the patrol route (cf. figure 7) is located in an area where many cells are ballistically threatening only if the semantics are not considered. Thus, Red attackers appear near the patrol route very likely and are neutralized with high probability. Thereby note that the patrol route has not been optimized but chosen quite arbitrarily by the user. Summarizing the ballistic semantics' effects, we claim that the impact is strongly related to the blue strategies and the amount of knowledge that has been used for their optimization. It is important that this optimization must consider the same semantic information as the simulation does. The reason for PfB to perform better without semantics and SoF to perform the same is random: The SoF patrol route has been chosen luckily and the far cells that are very unlikely to be reached in time from the base are not ballistically threatening when considering the semantics. But it is very important to have an accurate model in order to influence the strategy by optimization and simulation.

The answer to the second question is different. In all scenarios, the road semantics had no significant effect. For brevity, we do not show any results. We think the main reason was a realistic top-speed restriction on the road. Additionally, this restriction holds for red as well as for blue forces. Since both units had quite similar speeds (cf. table 1), they were restricted similarly on the roads and no influence of the success could be determined statistically. The road semantics might get interesting if it is considered for the setup of the strategy, e.g. during an optimization of patrol routes.

Comparison of Blue Strategies

During IDFW18, we also compared the different strategy options itself. Although this comparison was not our aim at this workshop, we can compare the strategies with tables 2, 3 and 4. Note that PfB is a sub-strategy of SoF as well as CE. Thus, we only have to compare the latter two strategies. CE is supported by the optimization module and the patrol route of the modeled SoF scenario has been chosen arbitrarily. The baseline strategy PfB is the worst one for Blue if we exclude the CE strategies with worst distributions. If we then distribute emplacements in an intelligent, i.e. optimized, manner, Blue's success rises as can be seen by the CE scenarios. Following strategy SoF with one patrol is nearly as successful (with the defined patrol route) as distributing 5 emplacements. This result basically confirms our results from IDFW18. The reasons for SoF's success are the following: On the one hand, the patrol in our scenario is in an area where attackers approach with high probability. Thus the time needed to disturb the Red is quite short. On the other hand, there exists a chance that Red senses Blue before starting to prepare its attack. Then it gets discouraged and flees without any attack. For these two reasons, SoF seems to be the best strategy for Blue. This superiority might even rise when the patrol routes are also optimized or more patrols are distributed.

Limitations

There are some limitations in our experiments. Concerning the optimization part, we can do a crossing of optimization and simulation runs. Although we optimized certain number of emplacements with and without considering the ballistic semantics, we only simulated the results of the former optimization. During the simulation, we then considered both ballistic grids again. In upcoming studies, we want to simulate the results of the latter optimization with both grids in order to check if a significant impact on Blue's success can be measured.

Another important limitation is that we are able to optimize the visibility of spotters, only. Additionally, we could define several optimization criterions, e.g. the reachability of possible attack points, in order to distribute QRFs outside the base also. One major challenge is the definition of the optimality if several criteria are considered. This is also subject to further research.

For an extensive comparison of the strategies, the following questions should be answered:

- How many emplacements/ patrols are needed to cover all cells?
- How can n emplacements/ patrols be distributed such that most cells are covered?
- What is a good ratio between covered cells and used emplacements/ patrols?
- How many emplacements/ patrols are needed to avoid any attack?
- How can n emplacements/ patrols be placed such that most attacks are avoided?
- What is a good ratio between avoided attacks and used emplacements/ patrols?

The first three questions can be answered for emplacements, but not for patrols since we currently are not able to optimize dynamic patrol routes. The last three questions could be answered by using simulation runs, but we have no hint for patrols to be defined without a corresponding optimization. Especially the third and sixth question is relevant since a notion of used resources and utility comes into account. The third question is basically answered for emplacements in figure 6.

The last limitation mentioned here is that only one specific Red behavior is modeled, which is also not evolving over time. We also always consider exactly a single attack and not multiple coordinated attacks which adapt to Blue's strategy. In order to challenge these questions at least semi-automated, we want to further extend our current approach with optimization techniques which are able to derive strategy settings automatically. Such a system could use evolutionary algorithms combined with data farming similar to Automated Red Teaming (ART) [3] and Automated Co-Evolution (ACE) [3].

CONCLUSIONS

During this study, we evaluated the influence of considering semantic attachments to our terrain data on Blue's success in terms of the regarded MoE PreventedShots. We wanted to check if it is worth to invest the effort to model or import such semantic attachments and model the reaction to it. We explicitly do not want to generate results like: "Blue's MoE decreases when considering semantics, so we do not consider it!" We claim that such a model is more realistic and thus more accurate. The question is if we want to invest the effort to improve such a model. As can be seen in the results of the optimization of emplacement positions, any information that is used to optimize some task must be based on an accurate model in order to generate correct results. For example, a spotter distribution generated on a ballistic grid considering semantics performs very well in a simulation considering ballistic semantics but not necessarily in a simulation without considering ballistic semantics.

We want to emphasize that the terrain data as well as the semantic attributes have not been modeled by us but are imported from a database of the German Armed Forces. We simply modeled the reaction to the attribute values, e.g. the top-speed restrictions of certain unit types on certain road types. Basically, we modeled effect of the semantics to the routing (road semantics) and to the ballistic threatening of cells (ballistic semantics).

The influence of the ballistic semantics is statistically significant if the information is used to set up the strategy. It is not significant and even arbitrary if we just evaluate a given scenario or strategy. The impact of road semantics is not significant in our study. This last result was a surprise.

Additionally, we wanted to check if our optimization according to the grid-visibility results in a significant improvement of Blue's success. During the investigation we learned that the latter question can be answered with a "yes", i.e. there is a significant improvement, if the model of the optimization is close enough to the model of the simulation.

As future work, we want to extend our approach as mentioned above. Additionally, we want to analyze the impact of the optimization, extend the system to be able to cope with several optimization criteria, develop an optimization for dynamic patrol routes and try to cope with coordinating and evolving red behavior.

REFERENCES

[3] D. Kunde, N. Mausberg, C. Mies, and B. Wolters, "Force Protection With ITSim: Base Protection Against Ballistic Weapons," in Scythe, Proc. and Bulletin Int. Data Farming Community, Issue 6, 2009, pp. 26-31.

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

[5] C.S. Choo, E.C. Ng, D. Ang and C.L. Chua, "Data farming in Singapore: a brief history," Proceedings of the 40th Conference on Winter Simulation, 2008, pp. 1448-1455

Team 6: Utility of Distillation Modeling for Countering IEDs

TEAM 6 MEMBERS

Byers, Ken
Lesnowicz, Ed
Meyer, Ted
McDonald, Mary
Upton, Steve
Naval Postgraduate School, US

Silwood, Narelle
Engleback, Nick
DTA, New Zealand

Middleton, Donna
Northrop Grumman, US

Malcolm, Paul
DSTO, Australia

Askman, Vic
Gunzelman, Karl
DeLange, Jon Paul
Lowry, Bruce
Other Organizations, US

INTRODUCTION

This Team participated in an ongoing study to examine the utility of distillation modeling in the Counter-IED (Improvised Explosive Devices) fight. In September, as part of that study, an Agent-Based Modeling workshop examining hard questions in the Counter-IED battle identified a set of problems that can be addressed by agent-based models and related methods such as data farming. The questions covered C-IED needs in various categories including: insurgent network evolution and adaptation; red-teaming and technical gaming; C-IED initiative assessment; and recidivism. This team reviewed the output from the September workshop on agent-based modeling, extracted problems that can feasibly be addressed in a rapid prototyping process, and began to design a software experiment that will address a selected problem.

We begin with some background information on our overall question, lay out our objectives and the effort for this IDFW 19 work, and conclude with a way ahead for follow-on work.

Background

In June 2009, JIEDDO began a study to examine the applicability and utility of agent-based modeling (ABM) and related techniques to its mission:

"The Joint Improvised Explosive Device Defeat Organization shall focus (lead, advocate, coordinate) all Department of Defense actions in support of Combatant Commanders' and their respective Joint Task Forces' efforts to defeat improvised explosive devices as weapons of strategic influence." JIEDDO Mission Statement, DoD Directive 2000.19E, February 14, 2006

The goal for this ongoing study is to address the following two questions:

- "Does ABM and related capabilities have applicability to JIEDDO questions and problems?"
- "How should ABM and related capabilities be applied to JIEDDO questions and problems?"

The answers to both "Does?" and "How?" may be different for different classes of questions or specific questions. As a result, part of the purpose of the ABM study is to provide analysts straightforward procedures to determine when ABMs might be applicable and what tools and experimental design is appropriate for their questions.

The ABM study encompasses two components or phases: 1) an educational component to gain an internal understanding and capability in these processes, and 2) an

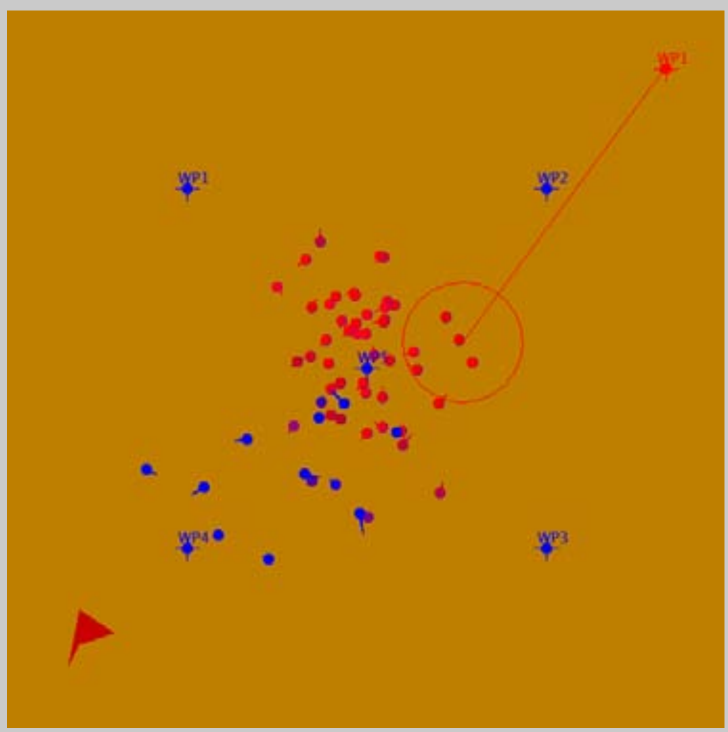


Figure 1 – Pythagoras "Peace" Scenario Spatial View

analytic component to apply these processes to questions of tactical, operational, and strategic interest to JIEDDO.

As one activity of Phase I or the Education Phase, an ABM Workshop was held in September, which focused on a set of JIEDDO questions and IED related problems and provided an opportunity for JIEDDO partners to give an overview of their current and planned work. The workshop consisted of five teams, producing 24 task plans that described the potential use of ABMs and other techniques to address relevant questions. Task plans included topics such as:

- Self-organizing graphs of data relationships
- Indirect Network Attack
- Identifying Important Link Layers for Impacting the Insurgent Networks in Afghanistan
- Define "High Value Individual"
- Emergence of an Insurgent Cell
- Insurgent Networks

The scope of related techniques of interest for addressing these study topics includes: agent-based models and modeling environments; social networking analysis tools; data farming tools; and analysis and visualization capabilities for model outputs.

As part of Phase II or the Analytic Phase of the ABM study, Team 6 at IDFW 19 began the task of selecting a subset of the questions of interest related to countering IEDs at the tactical, operational, and strategic levels and then identifying and applying appropriate ABMs and other relevant techniques to address those questions.

Objectives

Team 6's objectives for IDFW 19 were to:

- Examine Task Plans generated by the JIEDDO ABM Workshop
- Select potential candidate(s) for follow–up study and analysis
- Analyze and detail the question(s) being addressed
- Establish requirements for:
    - Modeling environment
    - Data requirements
    - Analysis tools
- Prepare plan ahead for activity between IDFW 19 and leading up to IDFW 20

IDFW19 EFFORT

The team began by examining the 24 task plans and selected a set of potential candidates to examine in further detail. These candidates fell into the category of "Attack the Network", one of three JIEDDO primary operations in the Counter-IED battle (the other two operations are "Defeat the Device" and "Train the Force"). These candidates are hard IED questions related to attacking the network and included 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 vs. networks that should be attritted vs. networks that should be co-opted?
- Will removing specific nodes destabilize a network? What are the 2nd and 3rd order effects? What are the potential unintended consequences?

Given that a network perspective is a primary component of the questions, the team decided to start looking at network analysis and visualization tools within the context of an abstracted "insurgent" scenario in the agent-based model Pythagoras. These tools may provide insights into JIEDDO topics of interest such as "Emergence of Insurgent Cells." We started with some simple networks to understand the issues and will progress to more complexity in follow-on work.

Why networks? Networks are useful representational schemes for understanding relations and interactions between agents (in our case, individuals) and events. Types of relations between two or more agents include such things as similarity (homophily) or spatial (distance) comparisons. Interactions might include events such as sensing or shooting, or other acts between agents. A network view, by using a different representation of relational data, has the potential of discovering underlying relationships that are hidden from

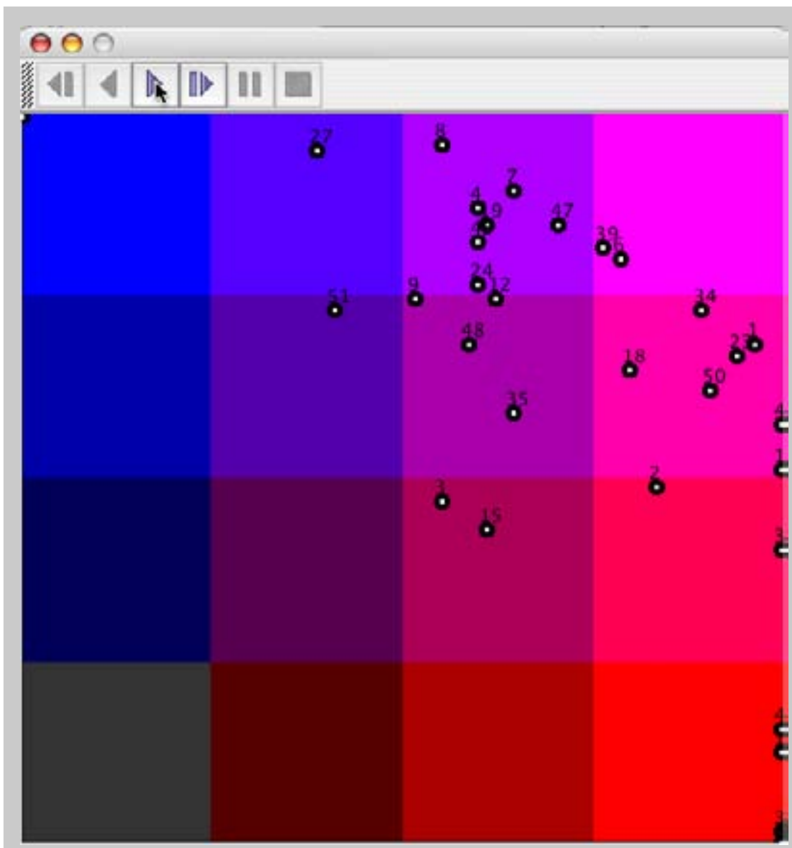


Figure 2 – "Peace" Scenario - Color Space View

other techniques, yet complements those techniques at the same time.

The next sections discuss the issue of modeling networks in an ABM, the Pythagoras scenario we used as an aid to understand network tools and analysis, and some results and surprises from our work.

Extracting Networks from ABMs

Since our ultimate goal is to understand and analyze insurgent networks, we wanted some means to model networks in an ABM implicitly and not explicitly, i.e., where the networks evolve dynamically based on agent attributes and behaviors and environmental characteristics and not fixed as input to the model. For example, both MANA and Pythagoras use a communications network to specify which agents send messages to other agents. However, those networks are fixed such that agent A cannot send a message to agent B if those links were not established as input to the model.

We then examined the use of an interaction-based network, where interactions between agents generate a dynamically evolving network. Pythagoras uses color attributes, i.e., red, green and blue (RGB) values for agents that are affected by changes in agent state, or a number of agent-agent interactions, such as shooting or communication. Pythagoras also allows the user to specify affiliation between agents by how "far" an agent is from another agent in color space, i.e., the distance (either Euclidean or Manhattan) between points based on the RGB values. The affiliations belong to four categories: Unit, Friend, Neutral and Enemy, so each agent has one of these affiliations for every other agent and these affiliations can be asymmetrical.

We next defined an affiliation network by using an RGB distance threshold such that if the RGB distance between two agents was less than the threshold, then a link was established. This network evolves over time based on agent RGB changes associated with interaction events, and was the focus of our network extraction efforts. We collected network data from a single run of a simple Pythagoras scenario, which we describe next, and visualized that data using several open-source tools.

Pythagoras "Peace" Scenario

The "Peace" scenario is provided as part of the Pythagoras distribution (we made some changes in the initial distribution of agents). The scenario is composed of a single Red Instigator, a population of 50 Purple Locals, and a small Blue force (10 agents). The Red Instigator is continually broadcasting pro-red/anti-blue messages to all the agents

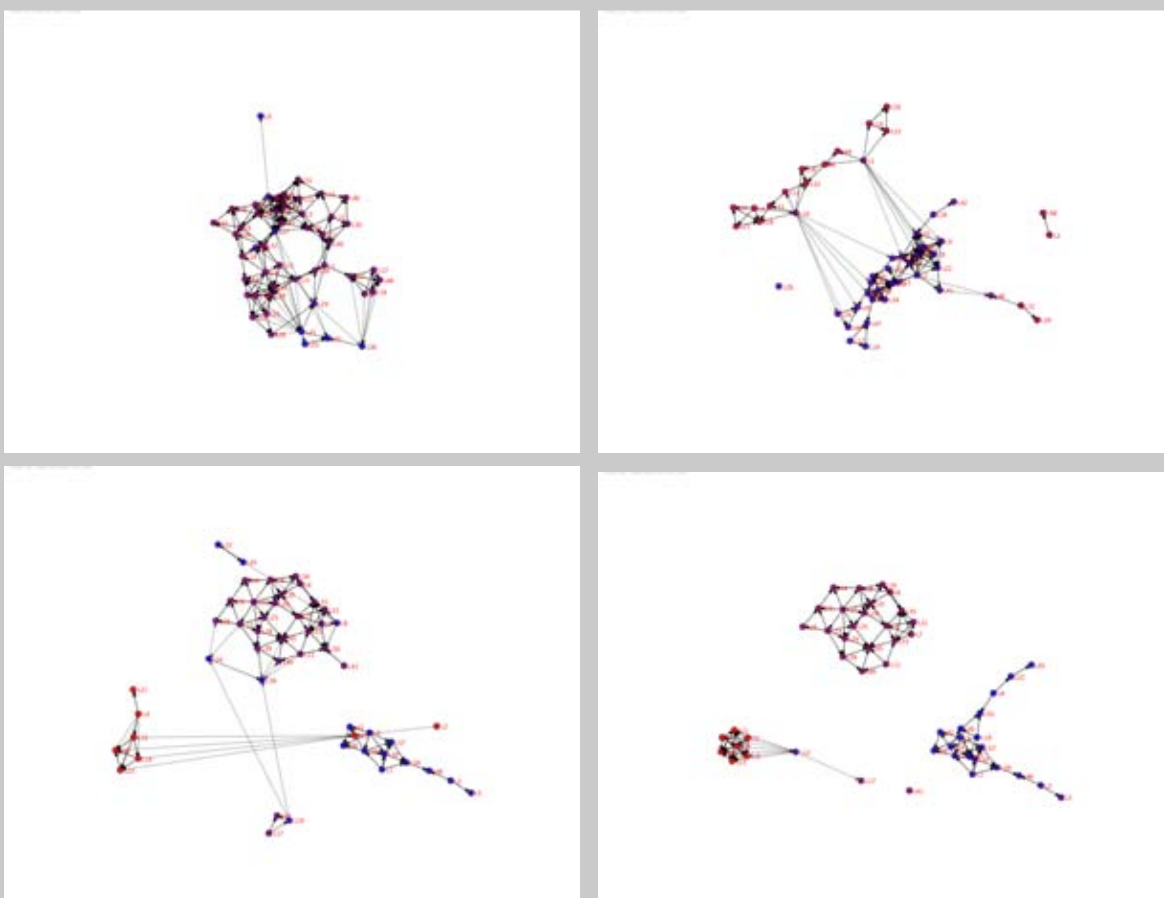


Figure 3 – "Peace" Scenario Network View

(although the messages currently have no effect on Blue), and the Blue force interacts "positively" with the Locals.

At each time step, the Red Instigator broadcasts a message that changes Purple Local agents within range of the broadcast by making them more "Red" and less "Blue". Concurrently, Blue agents move around and interact with Purple Locals within their range by making them more "Blue" and less "Red". These interactions have a random component, so that the competition between the Red and Blue messages affects each Local agent differently. A snapshot of one time step, showing a spatial view of the scenario, is depicted in Figure 1.

At each time step, we collected data representing each interaction that caused a color change of a Local agent, as well as the current color of all the Locals (because the Red and Blue agents didn't change color, our focus was on the evolution of the Local affiliation network).

In the next section, we describe two different visualizations of that data as well as some observations on how data farming might be applied to the visualization of network data, and not just for farming over input parameters of the model.

Results and Surprises

Figure 2 displays a "Color Space" view for one time step of the scenario. While not a network view, it does give a different perspective of the data, showing how each agent relates to the other agents based on their Red and Blue color attributes (Green was not used). The "closer" the agents are to other agents in this "space" indicates whether they are tied together in the network.

To derive a network view, we used the color distance and a threshold value to indicate which agents were affiliated with other agents. Figure 3 shows four affiliation networks for four separate snapshots in time. In each network, a node is a specific agent, and a link indicates that the two agents on either end of the link are within the threshold distance we specified. As the agents interact, their color changes resulting in the creation and deletion of links. And as those links evolve, other structures, such as cliques, form. Our goal is to understand the formation of these structures and how they are affected by changes in the input parameters of the model.

To display our results, we used the software package SoNIA (Social Network Image Animator - http://www.stanford.edu/group/sonia/), as well as some code we wrote during the workshop to translate Pythagoras data into a form suitable for display by SoNIA.

During our work, we encountered two surprises:

1. Extremely simple color distribution and interactions in Pythagoras lead to complex network interactions; and,
2. Data Farming over visualization/analysis parameters could provide additional insight.

WAY AHEAD

The plan for ongoing work between now and IDFW 20 and beyond will focus on modeling insurgent networks and continuing to look at network extraction and visualization tools and techniques, specifically:

1. Gain a better understanding of network analysis algorithms, animation, etc., especially as they relate to time-series or longitudinal data;
2. Expand the scenario to a more complex insurgent model, e.g., adapting a clique model previously implemented in Pythagoras;
3. Continue using Pythagoras, leveraging previous work and possibly running experiments on DOD HPC resources;
4. Examine the derivation of network statistics time-series and end-of-run MOEs as part of Data Farming analysis, optimizers, and ART (Automated Red Teaming) drivers; and
5. Examine methods of visualizing and comparing collectives of multiple networks, obtained by data farming across network parameter spaces.

Team 7: Social Simulations and Survey Data Focus Group

TEAM 7 LEAD

Steve Lieberman
Naval Postgraduate School, USA

INTRODUCTION

Understanding the social behavioral characteristics of the population in irregular warfare operations presents a major challenge to the defense modeling and simulation communities. However, a clear need has been identified for the development of models, methods, and tools to address the behavioral and social aspects of large human groups.

Objectives

The primary focus for this group was to provide an open forum for the discussion of cutting-edge approaches to modeling large collectives. Numerous questions were addressed regarding the goals, uses, and feasibility of different approaches to social simulation, from wholly conceptual to purely technical. Team 7 held three discussion sessions as follows:

- Session 1. The features of TRAC-Monterey's Cultural Geography Model, an agent-based model describing the social and behavioral patterns of civilian populations during COIN, were described.
- Session 2. Team members were tasked with the clarification of specific questions. The session addressed the nuances and limitations of current approaches to modeling large human groups, with discussion aimed at identifying data farming methodologies to address the limitations in what were most promising M&S methods.
- Session 3. An overview of team 7 was provided.

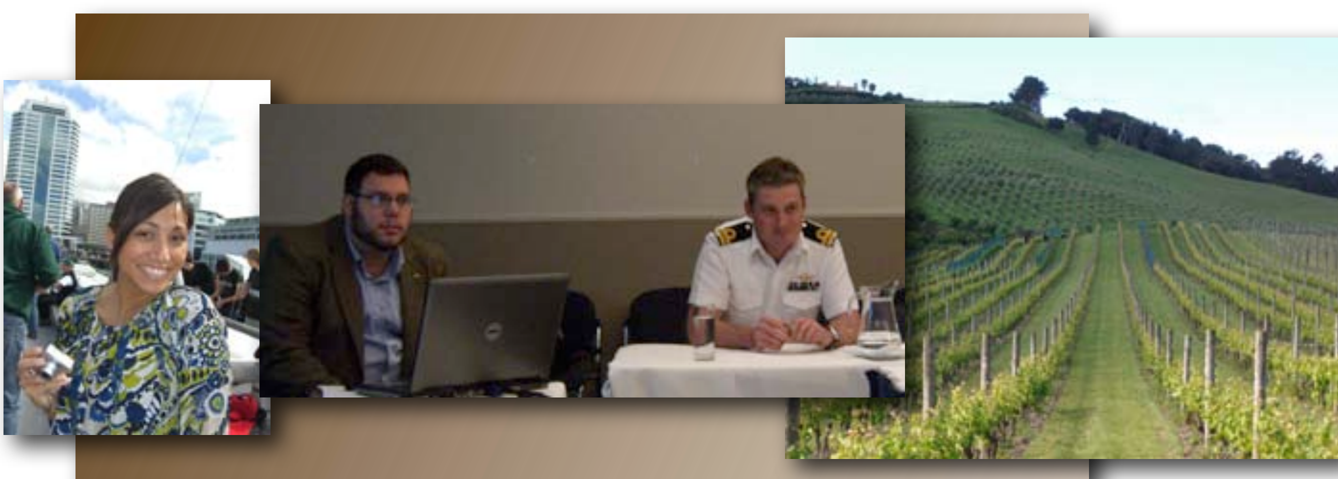
Through the discussion sessions the team had the following goals:

- From a security and defense perspective identify the most important differences between modeling Large Groups vs. Small Groups.
- Identify the goals of modeling societies.
- Identify the data that is required for social simulation.
- On a conceptual and technical level, find out how to deal with the conflation of belief and behavior change.
- Find the most effective ways to communicate model and simulation output to analysts or decision makers.

SUMMARY AND WAY AHEAD

Based on the Team 7 discussions resulted in the following food for thought as we move forward:

- Agent-Based Models/Multi-Agent Systems provide a viable body of simulation literature, including compelling ideas from individual (cognitive psych), team-based (social psychology) and organizational (network) modeling and simulation.
- Combining agent-based modeling and network modeling techniques can produce authentic representations of dynamic social structures that are especially important for course of action analyses
- Widespread interest in social simulation, within both the defense communities, and academic communities, has underscored the need to cultivate new research communities for modeling societies.



Team 8: Data Farming in Support of NATO

TEAM 8

Dr. Gary Horne, Chair
USA

LTC Stephan Seichter, Co-chair
DEU

SUMMARY

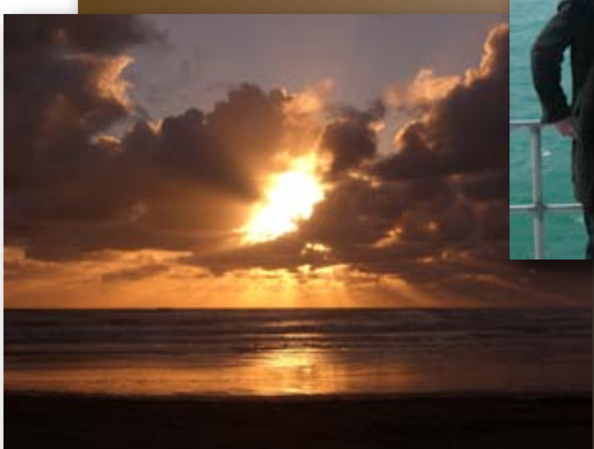
The NATO Modeling and Simulation Group has approved an exploratory team to examine the veracity of forming a task group to examine data farming in support of NATO. This exploratory team, called ET-029, continued its activity at IDFW 19, convening focus groups to further this examination.

Given the following basic points, at the NATO Modeling and Simulation Group meeting in October the exploratory team received support from 10 nations for consideration as a task group,.

- The nature of scenarios that NATO forces are faced with in today's world are uncertain and complex.
- Data farming combines rapid prototyping of agent-based and other models with the exploratory power of advanced computing to rapidly generate insight into questions.
- Data farming allows the decision maker to more fully understand the landscape of possibilities and also allows for the discovery of outliers.
- Six realms of data farming are model development, high performance computing, visualization of large simulation data output, rapid prototyping of scenarios, parameter space exploration through efficient design of experiments, and collaborative processes.

This task group, pending final approval, will be called MSG-088 and will hold its first official meeting in May 2010 in Paris. Here at IDFW 19, possible support areas of work were discussed, including hardware, overview of models, example scenarios, efficient designs, visualization, analysis tools, evolutionary algorithms, effects based planning, and decision support. Application areas thought to be worthwhile include winning hearts and minds and the cascading consequences of actions in that area, command and control and the interface challenges of the multi-national forces of NATO, peace support operations, urban operations, anti-piracy operations, training, and defeating improvised explosive devices.

Team 8 concluded with plans to contribute to further discussions of potential MSG-088 efforts. These plans include another ET-029 meeting to be held following IDFW 20 on Friday 26 March 2010 in Monterey, California, USA. Please contact gehorne@nps.edu if you are interested in participating.



International Data Farming Workshop 20

When: 22 - 25 March 2010

Where: Naval Postgraduate School, Monterey, California

Data farming is a capability that makes use of high performance computing to run models many times. This capability gives modelers and analysts an enhanced ability to discover trends and outliers in results, do sensitivity studies, verify and validate over extended ranges of input parameters, and consider modeling and analyzing non-linear phenomena with characteristics that cannot be precisely defined. The International Data Farming Workshops have been a forum for the exploration of important questions for many years. The cores of the workshops are teams that are formed around these questions. These multi-disciplinary teams use data farming practices including simulations, such as agent based models, rapid model prototyping, high performance computing, and data analysis, state-of-the-art design of experiments, parameter space exploration, and collaborative environments. The first workshop took place in Maui in 1999 and the workshop coming up in Monterey next month is number 20. We now have 17 teams lined up for IDFW 20 and whether you are a newcomer to the workshops or have participated in the past, we welcome you to participate!

IDFW 19 2009 Tentative Agenda

- **Sunday, March 21:** Opening reception and dinner
- **Monday, March 22:** Opening briefs and team poster sessions in the morning, then begin work in teams
- **Tuesday - Wednesday, March 23 - 24:** Work in teams (optional plenary sessions in the mornings)
- **Thursday, March 25:** Outbriefs and Closing Ceremony in the afternoon

Call for Team Leaders / Plenary Speakers:

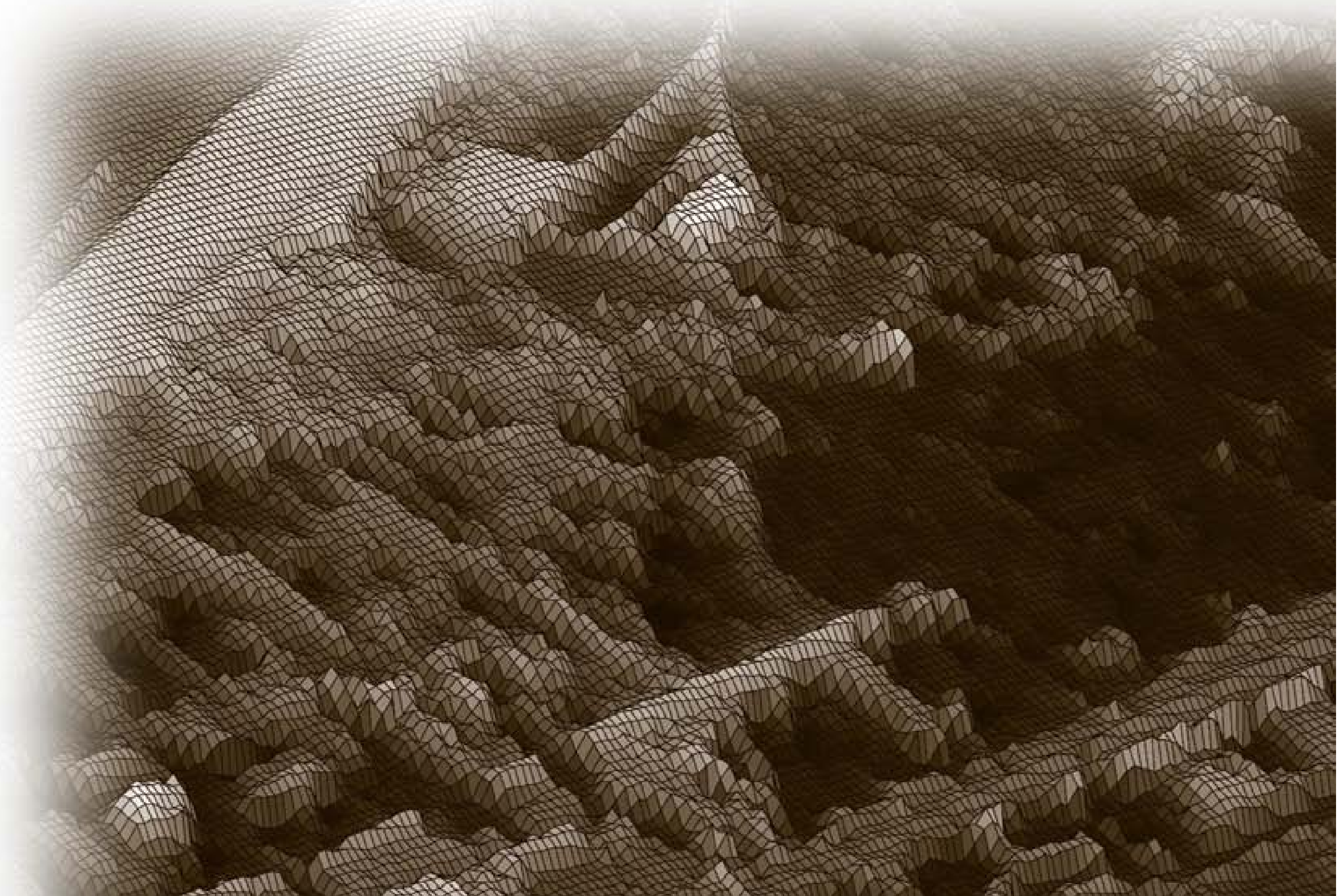
Please email gehorne@nps.edu with your choice of teams and if you want to lead a team or present a plenary briefing.

Theme: Enrichment

International Data Farming Workshop 20

March 22-25, 2010

Monterey, California



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 7 - Workshop 19