Data Farming Around the World Overview

Gary E. Horne, Ph.D.

SEED Center for Data Farming
Naval Post Graduate School
Monterey California, U.S.A.
e-mail: gehorne@nps.edu

Klaus-Peter Schwierz, Ph.D.

System Design Center Germany
European Aeronautic Defense and Space Company (EADS)
Friedrichshafen, Germany.
e-mail: Klaus-Peter.Schwierz@EADS.com

Dr. Gary E. Horne is a research professor at the Naval Postgraduate School (NPS) in Monterey, California, in the USA. He holds a B.S in Mathematics and an M.Ed. in Education from the University of Maryland and a D.Sc. in Operations Research from the George Washington University. He is the originator of the Data Farming concept and the Executive Director of the SEED (Simulation Experiments and Efficient Designs) Center for Data Farming at NPS.

Dr. Klaus-Peter Schwierz is a senior manager in the System Design Center of EADS DS (European Aeronautics Defense and Space Company, Defense & Security) in Germany. After education in Mathematics and Physics he holds a Diploma in Mathematics and a D.Sc. in Natural Sciences / Mathematics from the Albert Ludwigs University in Freiburg, Germany. His business position is characterized as a consultant for system and system of systems decisions with the help of simulation. He and many of his colleagues are members of the Data Farming Community, attracted by the methodology and are promoting the Data Farming Concept in Germany.

ABSTRACT

Data Farming, a question based process, combines the rapid prototyping capability inherent in a certain class of highly abstract, fast running simulation models with the exploratory power of high performance computing to rapidly generate insight into questions. The Data Farming process focuses on a more complete landscape of possible system responses, rather than attempting to pinpoint an answer. The focus is on a continuum of solutions, looking for (unknown effects) and interrelations, the analysis of a variety of possible progressions and a consequent application of optimization theory. Thus Data Farming allows decision makers to more fully understand the landscape of possibilities and also consider outliers that may be discovered. Over the past decade, an international community has formed around these ideas dealing with modeling of non-linearity, intangibles and co-evolution. In the fall of 2008, International Data Farming Workshop 17 took place in Garmisch Partenkirchen, Germany and workshop number 18 was held in spring 2009 in Monterey, California, USA, where the

SEED Center for Data Farming at NPS is the home base of Data Farming. In addition to a summary of the two last workshops, this paper will present an overview of the process that has developed to include the development of both methods and applications in the International Data Farming Community.

ACKNOWLEDGEMENTS

Instantiating the idea of Data Farming in the real world of modeling and high performance computing has benefited from the patience, effort and interest of a large number of participants from around the world. The authors would like to acknowledge the integrity, audacity, and humility of all of the collaborators in the International Data Farming Community: the modelers, software developers, subject matter experts, analysts, administrative aids, students, teachers, and decision makers. We are especially grateful for the insight, interest, and time provided by men and women in military service aro und the world. But most of all we want to express our gratitude to our contractor theGerman Armed Forces who have funded this important analysis

INTRODUCTION

The ideas behind Data Farming were introduced by Dr. Horne to the defense community in 1997 (Horne, 1997). In 1998 Dr. Brandstein, chief scientist at the USMC at that time, and Dr. Horne (Brandstein and Horne, 1998) presented during the Cornwallis Group II: Analysis for and of the Resolution of Conflict the ideas under the title “Asymmetric Evolutions and Revolutions in Military Analysis.” In concert with the combination of agent-based models with high performance computing that was the start of Project Albert.

Project Albert was a congressionally funded modeling and simulation initiative of the United States Marine Corps (USMC) motivated by the fact that complex adaptive systems are pervasive in USMC operations. The models had to deal with non-linearity, adaption, co-evolution, intangibles and emergence, just to mention a few essential factors appearing in these operations. Agent Based Models are one expression of complex adaptive systems. The philosophy of Project Albert was to pair these simple, efficient, abstract models with high performance computing to explore large design spaces. When these models and high performance computing are combined with efficient experimental designs developed in work pioneered at NPS (e.g. see Sanchez and Lucas, 2002), a huge sample space can be explored very rapidly. And when rapid prototyping capabilities and collaborative environments are introduced into the Data Farming process, progress on questions, even long-standing and difficult questions involving many interacting variables, is possible.

Project Albert used what are referred to as agent-based distillation models. These are a type of computer simulation which attempts to model the critical factors of interest in combat without explicitly modeling all of the physical details. Some of the models used in Project Albert were MANA, PAX, and Pythagoras, all agent-based models, although the methods developed can be applied using any type of simulation model. These models continue to be developed and recent updates are described in Lauren (2007), Lampe (2007), and Henscheid (2007) respectively. In addition these agent-based models are small and abstract and can easily be run many times to test a variety of parameter values and get an idea of the landscape of possibilities. The term distillation is added, because the intent is to distill the question at hand and down into as simple a representation as possible. Also, models used in Project Albert were specifically developed and used because the capability to rapidly prototype scenarios is also very important in this process.

Although Project Albert was a US-sponsored effort, it had a strong spirit of international collaboration which made a great deal of cooperative effort among researchers around the world possible. For example, the three models mentioned in the paragraph above were sponsored by MoD´s in New Zealand, Germany, and the US respectively. Due to the fact, that Project Albert was question-based also allowed practitioners from around the world to rally around the developing Data Farming methodologies because of the impact it had upon their present shared application interests.

Because many of the questions of interest have wide applicability, the work teams at international workshops typically consist of representatives from minimum three and sometimes up to six different countries and the workshops overall have usually consisted of about 8 to 12 teams ending in more than 120 applications of the models to specific mission related questions. The first international workshop was organized in 1999 and 12 workshops were held under the auspices of Project Albert which ended in September 2006. Since that time the international Data Farming community has continued cooperative efforts and the most recent workshops, the 17th and 18th workshops were held in fall 2008 and in spring 2009. Our presentation here will highlight these most recent workshops, but details on the other workshops and other Data Farming efforts can be found at http://harvest.nps.edu, the website of the SEED (Simulation Experiments and Efficient Designs) Center for Data Farming at the United States Naval Postgraduate School.

DATA FARMING

OVERVIEW

Data Farming combines efficient experiment design, the rapid prototyping of agent-based distillations with the exploratory power of high performance computing to rapidly generate insight into military questions. Data Farming focuses on a more complete landscape of possible system responses and progressions, rather than attempting to pinpoint an answer. This "big picture" solution landscape is an invaluable aid to the decision maker in light of the complex nature of the modern battle space. And while there is no such thing as an optimal decision in a system where the enemy has a role, Data Farming allows the decision maker to more fully understand the landscape of possibilities and thereby make more informed decisions. Data Farming also allows for the discovery of outliers that may lead to findings that allow decision makers to no longer be surprised by surprise.

The simulations that defense analysts use are often large and complex. An evaluation of complete landscapes is extremely time consuming, sometimes not even possible. Also, even the smaller more abstract agent-based distillations referred to above can have many parameters that are potentially significant and that could take on many values. And response surfaces can be highly non-linear. Thus, even with high performance computing and the small models used in Data Farming, gridded designs, where every value is simulated, are unwieldy.

Thus, the statistical approach, using efficient experimental designs is essential and work in this area has been performed at the Naval Postgraduate School (NPS) in Monterey, California and NPS researchers have collaborated with others worldwide as well (see Kleijman, Sanchez, Lucas, and Cioppa, 2005). Data Farming continues to evolve from initial Project Albert efforts (Hoffman and Horne, 1998) to the work documented in the latest edition of the *Scythe* (Horne and Meyer, 2009). This proceedings and bulletin of the International Data Farming Community contains the proceedings of the International Data Farming Workshops that have taken place since Project Albert ended and is put out by the SEED Center for Data Farming at NPS.

QUESTION-BASED

Over the past few years several articles have captured the fundamentals of Data Farming (e.g. Horne and Meyer, 2005 and Lawler, 2005), but the key is the question basis. Over the past decade and as we shall present in the next section, over 120 international work teams have formed around questions at International Data Farming Workshops. On the other hand, at the Naval Postgraduate School over 60 theses, many by international students, have been completed which have used data farming.

These types of questions can never have precisely defined initial conditions and a complete set of algorithms that describe the system being considered. These questions address open systems that defy prediction. Data Farming is used to provide insight that can be used by decision-makers. To accomplish this challenging task, Data Farming relies upon two basic ideas:

1. Use high performance computing (HPC) to execute models many times over varied initial conditions to gain understanding of the possible outliers, trends, and distribution of results, and

2. Develop models, called distillations, that are focused to specifically address the question.

The discovery of surprises (both positive and negative) and potential options are made possible by Data Farming, by providing the ability to process large parameter spaces.

ITERATIVE PROCESS

Data Farming is a collaborative and iterative team process (Horne and Meyer 2004b). Figure 1 describes the Data Farming process as a set of imbedded loops of iteration. This process normally requires input and participation by subject matter experts, modelers, analysts, and decision-makers.

The “Scenario Building” loop shown on the left side of the Figure involves developing and honing a model that adequately represents the system that addresses the question being asked by the decision-maker. This is an iterative process that often requires honing the question as well.

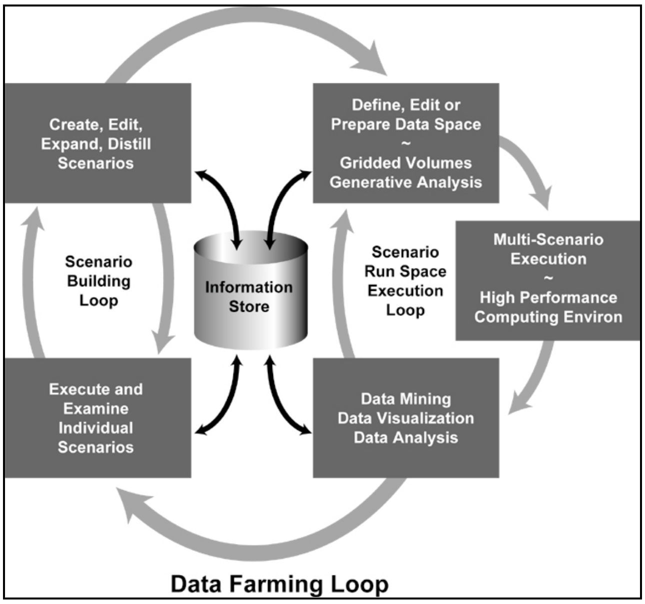

Figure 1: Data Farming is an Iterative Process.

The "Scenario Run Space Execution" loop shown in Figure 1 is entered once the *base* ***case*** of the scenario is complete. In this loop the team defines an *experiment* which determines which scenario input parameters should be examined and what processes should be used to vary them. Here the team is exploring the possible variations (or *excursions* of the base case) in the initial conditions of the scenario. Specifically those parameters that address the question being posed are considered.

The defined experiment is used to guide the execution of many runs of the model in the HPC environment. Each run produces output which is collected by the Data Farming system and provided as output to analysis capabilities. After analysis of the results, the team (or an algorithm) may decide to adjust or produce a new study or adjust the model to more adequately address the question. This process continues until insight related to the decision-maker's question has been gained.

WORLD WIDE PARTNERS

The first international workshop took place in 1999. The first 4 workshops were methodology driven, dealing with Complex Adaptive Systems Modeling and the agent based representation, with Statistical Experiment Design and Experiment Evaluation. The subsequent 14 workshops were application driven, contributions to the overall advancement of Data Farming takes place in the development of simulation models, scenarios within the models, and computer clusters to run the models audacious numbers of times. The real work is in making progress on important questions and the real secret is the combination of military subject matter experts and highly knowledgeable and multi-disciplinary scientists. This special mix of personnel has been the hallmark of the international workshops. It has been a dynamic combination to have Data Farming work teams headed up by a person who really knows and cares about the question (e.g. a military officer who knows that the answers may have an impact on both mission success and lowering casualties) and supported by men and women with technical prowess who can leverage the tools available. The international community forming around the ideas of Data Farming is shown in Figure 2. The workshops took place in following cities: Kihei, Auckland, Cairns, Überlingen, Monterey, Quantico, Singapore, Wellington, Stockholm, Honolulu, Boppard, Den Hague, and Garmisch Partenkirchen.

Figure 2: The International Data Farming Community.

Now we will now provide some details regarding how German involvement increased from the participation in the early workshops when the methodologies were in the beginning stages of development. After Project Albert started the emphasis on using Data Farming tools within combat situations and the simulation tools were developed to represent these situations, the German Delegation carried in questions regarding peace support operations. Human factors modeling and the influences of intangibles are becoming more and more essential in this question area. To simulate the non-attrition based parts in peace support operations the model PAX (after the Roman goddess of peace) was developed in Germany and released to the International Data Farming Community. The contributions led to a broad acceptance of Data Farming in the German modeling and simulation community.

All German applications had the clear goal: Not to replace the classical Modeling and Simulation tools by new ones but to apply both methods in an "operational synthesis" (see Brandstein, 1999). The application of complex adaptive systems theory with the modeling following the agent based paradigm had the goal to explore the wide field of non linearity, of co-evolution and intangibles. Results were a continuum of solutions in the sense of optimization theory with the relevant tools for a statistical experimental design and the semi-automated evaluation techniques directing the user to unknown effects, or "surprises" and interrelations in the analysis of a variety of possible progressions. Figures 3 and 4 show a comparison of "What we did" and "What we got" as German delegations by joining the Data Farming Community.

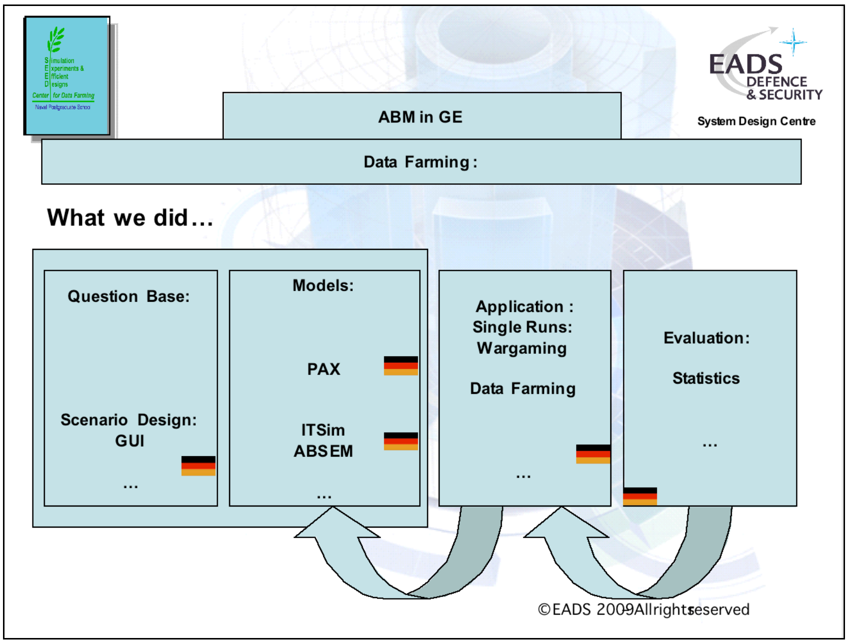

Figure 3: Benefits obtained by joining the Data Farming Community.

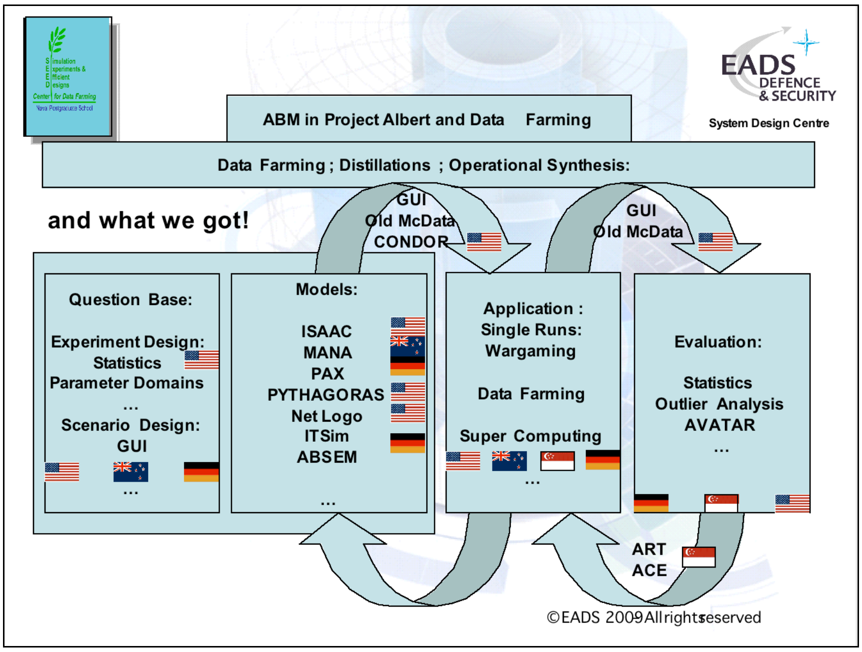

Figure 4: Benefits obtained by joining the Data Farming Community.

For the other attending nations Figures comparable to Figures 3 and 4 can be drawn. The international community drove the German mission through other methodological developments such as Nearly Orthogonal Latin Hypercube experimental designs (e.g. see Cioppa, 2002), model developments (MANA, Pythagoras, PAX, etc.), application of agent based-development environments (NetLogo, REPAST, MASON, etc.) and through evaluation and analysis tools of many types (e.g. see Upton, 2004). In the international workshops the availability of the experts and the free and open information sharing led to a big success in Germany and all other attending nations in the application of the tools (Schweirz, 2008).

Other Data Farming efforts around the world are documented in a variety of places. The beginning of development in the United States is documented in *Maneuver Warfare Science 1998* (Hoffman and Horne, 1998). And additional volumes of Maneuver Warfare Science from 2001, 2002, and 2003 contain contributions from the US as well as Sweden, New Zealand, Australia, and Singapore. Also, the book *It's Alive* (Meyer and Davis, 2003) contains a chapter describing some of the initial USMC efforts. Many presentations involving Data Farming have also been made at INFORMS meetings (e.g. Horne and Meyer, 2004a) and MORS Symposia (e.g. McDonald, 2008) over the past decade and Winter Simulation Sessions on Data Farming were held in 2004, 2005, and this year. Finally, the *Scythe* is a regular publication from the SEED Center for Data Farming that documents workshop proceedings.

INTERNATIONAL WORKSHOP TOPIC SUMMARY

Through 14 international workshops, which were application driven, we have had over 120 work teams in a variety of areas. Of course some of those work teams have continued from workshop to workshop. For example the Peace Support Operations team which has had representation over the years from 7 different countries started at workshop 5 and continues to the present. These 120+ work teams do fall into areas, or themes, which include: Joint and Combined Operations (e.g. C4ISR Operations, Network Centric Warfare, Networked Fires, and Future Combat Missions), Urban Operations, Combat Support (e.g. UAV Operations, Robotics, Logistics, and Combat ID), Peace Support Operations, the Global War on Terrorism, Homeland Defense, and Disaster Relief. Essential is, that in every workshop there were working groups dealing with: Urban Operations, Peace Support Operations and C4ISR Operations. Starting in workshop 8 a sensor – effector modeling team attended all following workshops and in 7 of the workshops a combat identification team worked with the support of 6 nations. Other work teams have looked into continuous specification support of modeling such as efficient designs, new models, model improvements, automated red teaming, and automated co evolution.

INTERNATIONAL WORKSHOPS IN FALL 2008 AND SPRING 2009

Two Workshops were held in fall 2008 and in spring 2009. IDFW 17 under the title "Discovering Surprises" was held in Garmisch Partenkirchen, Germany 21 to 26 September 2008 and International Data Farming Workshop (IDFW) 18 under the title "Dynamic Truths" was held in Monterey, California, USA from 22 to 27 March 2009. Here we list the titles the working groups of both workshops to give a flavor of the breadth of topics explored at an IDFW (Meyer and Horn, 2008b).

IDFW 17: DISCOVERING SURPRISES

1. Team 1 - PAX Refugee Camp Scenario, Phase I: Model Calibration. Team Leader: Germany.

2. Team 2 - Exploring Peace Support Operations using PSOM. Team Leader: USA.
3. Team 3 - Modeling the Effects of Human Intangibles. Team Leader: Singapore
4. Team 4 - Evaluation of Electro-Optical Sensor Systems in Network Centric Operations using ABSEM. Team Leader: Germany.
5. Team 5 - Effects of Terrain on Systems of Systems. Team Leader: Singapore.
6. Team 6 - Exploring the Design Space of Command and Control Capability Evaluation Strategies. Team Leader: USA.
7. Team 7 - Growing a PMESII Community of Interest. Team Leader: USA.
8. Team 8 - Data Farming the Basic Immune Simulator. Team Leader: USA.
9. Team 9 - Exploring Communications in an Urban Environment. Team Leader: USA.

IDFW 18: DYNAMIC TRUTHS

1. Team 1 - Maritime Protection Offshore. Team Leader: New Zealand.
2. Team 2 - Data Farming Unmanned Systems Used in Casualty Evacuation. Team Leader: USA.
3. Team 3 - Total Life Cycle Management (TLCM). Team Leader: USA.
4. Team 4 - Evaluation of Sensor and Effecter Systems in Network Centric Operations using ABSEM 0.2. Team Leader: Germany.
5. Team 5 - PAX Modeling. Team Leader: Germany.
6. Team 6 - Joint Mission Effectiveness support using Data Farming (JMEDF) in Netcentric Systems Test Planning. Team Leader: USA.
7. Team 7 - Investigating the use of simulation tools for mass casualty disaster response. Team Leader: USA
8. Team 8 - Simulation of Unmanned Aerial Vehicles to Enhance Border Security. Team Leader: Turkey.
9. Team 9 - Logistics Battle Command. Team Leader: USA.
10. Team 10 - Peace Support Operations Model. Team Leader: USA.
11. Team 11 - NATO Frigate Defense Effectiveness vs. Small Vessels in Open Waters. Team Leader: Germany
12. Team 12 - Representing Cultural Geography in Stability Operations. Team Leader: USA.

IDFW 17 DETAILED EXAMPLES

In order to get a better understanding of the type and quality of work which is being performed at these International Data Farming Workshops two following examples of the areas of study are presented:

- Team 1: PAX Refuge Camp Scenario Phase I: Model Calibration in Preparation of Validation (Seichter, S. and Schwarz, G., 2009).
- Team 4: Evaluation of Electro-optical Sensor Systems in Network Centric Operations using ABSEM 0.1 (Haymann, K. and Wagner, G., 2009)

PAX REFUGE CAMP SCENARIO PHASE I: MODEL CALIBRATION IN PREPARATION OF VALIDATION (SEICHTER, S. AND SCHWARZ, G., 2009)

The Centre for Transformation of the German Armed Forces has initiated a series of studies to explore the requirements for the implementation of Modelling & Simulation for an effective support of experiments in the Concept Development & Experimentation process (CD&E). One study examines possibilities to perform analyses using PAX in the context of a multinational Peace Support Operation including Humanitarian Assistance dealing with disaster caused refugee movements that require military action. Military forces are tasked to assist in building and operating refugee camps, especially to ensure order and security.

Basic questions under examination in this new scenario may be:

- How to equitably distribute goods such as water and food among the different refugee groups with individual attitudes for aggression and fear?
- Are there dominant groups and, if so, why?
- Are there any groups or individuals unable to receive the needed goods?
- How can military forces successfully distribute food even-handedly?
- How best to prevent aggressive escalation in the refugee camp?
- Is separation a successful way?
- Which strategy should be chosen by the military forces - ranging from de-escalation to deterrence?

The development and implementation of new scenarios requires the calibration of the model in order to achieve a sufficiently high degree of validity for that specific purpose. Therefore the team's primary goal during the IDFW 17 was to develop and refine appropriate calibration methods. This included the identification of suitable Measures of Effectiveness (MoEs) and the necessary set of parameters to vary during calibration as well as examining different calibration methods including large experiments with various Designs of Experiment (DoE), optimization using Automated Red Teaming (ART) functionalities and a focused mini Experiment.

Brief Scenario Description

The applied scenario models the distribution of aid packages in a refugee camp supervised by soldiers. Three major groups of refugees are present within the camp. These groups are characterized either as "normal," "cautious," or as "aggressive" civilians.

The actual questions to be analyzed are for example:

- Identification of equitable and evenhanded distribution strategies with respect to the different refugee groups and their individual emotional attitudes like aggression and fear.
- Identification of groups or individuals unable to receive the needed goods.
- Effectiveness of military strategies, ranging from de-escalation to deterrence in the given context.

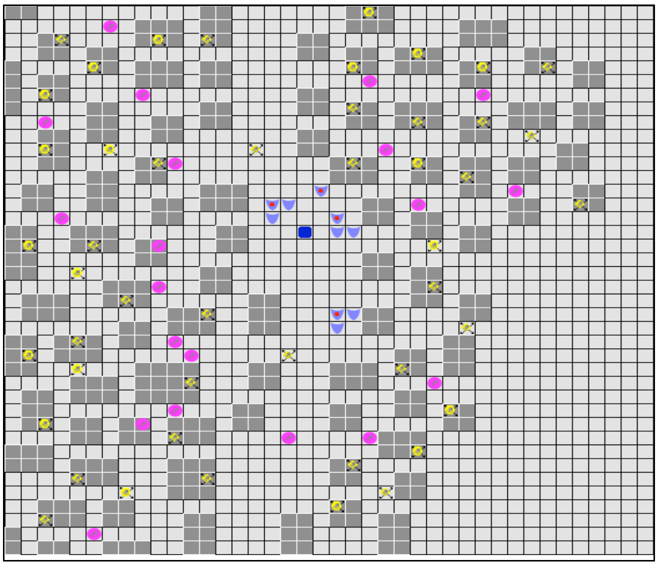

Figure 5: Refugee camp (“macro scenario”) in PAX.

Approach for Calibration

Preceding the calibration efforts, important MoEs to determine the quality of calibration had to be ascertained. The related discussion, involving "subject matter expertise" on refugee camp management, lead to the identification of the following MoEs reflecting the most important indicators of agent behavior:

- Total number of aid packages distributed.
- Number of aid packages distributed to {aggressive, normal, cautious} civilians.
- Aggregated escalation.
- Number of {attack, threaten} actions performed by civilians,
- Number of pacifying actions performed by soldiers.

Taking into consideration the structure and interplay between PAX model parameters, a minimal but sufficient parameter subset were determined to be vary in order to calibrate the model to the MoEs and to have a manageable experiment size besides (Table 1, values in the middle column). Furthermore, an initial process for calibration was developed combining *large scale calibration*, *micro scenario analysis* and *ART optimization*.

The term *large scale calibration* defines a cyclic procedure with appropriate experimental designs that iteratively adjusts the respective parameter ranges until they can be evaluated as satisfying for the given setting.

The *micro scenario analysis* looks at an appropriately downscaled scenario with the goal of identifying *realistic ranges* for the parameter(s) subject to calibration. Having evaluated the results to the desired degree, the applicability of the effects in a macro scenario has to be proofed.

Optimization using ART seeks to optimize selected MoE(s) towards desired, better more realistic and feasible target values.

Parameter to be calibrated	Range/Val Exp #1	Range/Val Exp #2
Aggressive group Dog Factor	[0.5;3.0]	[0.5;**5.0**]
Normal group Dog Factor	[0.5;3.0]	[**2.0**]
Cautious group Dog Factor	[0.5;3.0]	[**2.0**]
Aggressive group –	[70;100]	[**90**;100]
Normal group Need	[70;100]	[**80**;**90**]
Cautious group Need	[70;100]	[70;**80**]
Aggressive group PC_Anger	[0;50]	[**10**;50]
Normal group PC_Anger	[30;100]	[30;**80**]
Cautious group PC_Anger	[30;100]	[**50**;100]
Aggressive group PC_Fear	[30;100]	[30;100]
Normal group PC_Fear	[30;100]	[30;**80**]
Cautious group PC_Fear	[0;50]	[**10**;50]
Aggr. Group RfA	[30;100]	[30;100]
Norm. group RfA	[0;Aggr.]	[**10**;Aggr.]
Caut. group RfA	[0;Norm]	[**10**;Norm]
Soldiers' stress factor	[0;100]	[0;**5**]
Soldiers' ideal stress value	[0;100]	[**20**;**40**]
Escalation until arrest threshold	[10;200]	[10;**100**]

Table 1: NOLH experiment (1+2) parameters for calibration and value ranges explored.

Large Scale Calibration and Micro Scenario Analysis

Large Scale Calibration and Micro Scenario Analysis is covered in reference (Seichter, S. and Schwarz, G., 2009) and have been skipped in this abbreviated version of the PAX Refuge Camp Scenario Phase I: Model Calibration in Preparation of Validation.

Optimization Using the ART Tool

The goal of applying ART's (Automated Red Teaming, DSO National Laboratories, Singapore) optimization functionalities was to optimize selected MoEs towards subject matter experts' expected outcomes. It should be noted that ART was not used for its original purpose of red teaming. We rather used its implemented optimization algorithms on selected MoEs in order to find respective feasible input parameter settings.

Table ***1*** shows the experiment setup with the large-scale initial settings in column 1.

Column 2 lists the respective parameter ranges that lead to the target values in the selected MoEs as shown below. The considered MoEs comprised the number of packages distributed to each refugee group because there were pretty clear subject matter experts' expectations for these outcomes. The target values that ART was given to optimize towards are displayed in

Table 2, compared to the values achieved with ART.

Parameters submitted to ART	Parameter Values in PAX	
	Search Range	"Optimized Range"
Aggressives' Pers. Const. Anger	10-50	11-29
Aggressives' Readiness for Aggr.	30-100	57-99
Cautious' Pers. Const. Fear	10-50	19-44
Soldiers' Stress Factor	0-5	0.7-4.3
Aggressives' Pers. Const. Fear	30-100	36-77
Aggressives' Get many packages	90-100	91-96
Normals' Pers. Const. Anger	30-80	35-77
Cautious' Pers. Const. Anger	50-100	56-87

Table 1: ART experiment search and optimized ranges.

Measure of Effectiveness	Target Value	Closest Value Achieved
# Packages distributed to civilians of aggressive group	20	22
# Packages distributed to civilians of cautious group	35	26
# Packages distributed to civilians of normal group	60	61

Table 2: Desired values towards which ART was set to optimize the scenario.

Figure 6 displays these three MoEs (average over 10 replications) for all scenarios that ART produced after the 4th generation of the SPEA2 evolutionary algorithm that was used, which had been configured to use an initial population size of 60, an archive size of 30 and a maximum of 8 generations.

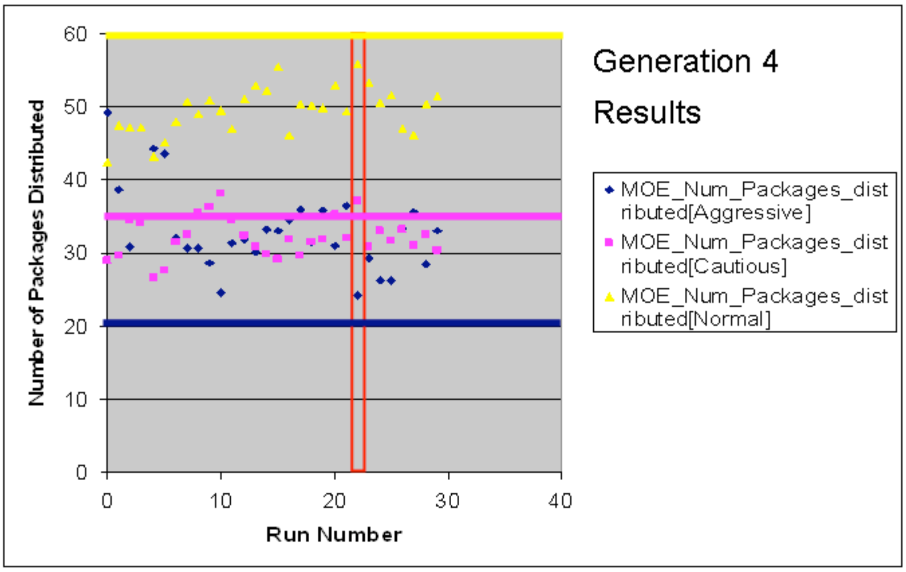

Figure 6: MoEs of the scenarios calibrated using ART with PAX and the refugee camp scenario.

Due to time limitations, only four generations could be performed and analyzed. But the results were already convincing. Figure 14 lists the results, the average MoEs were quite close to the target values. It is to be expected that the remaining generations of evolved scenarios will yield even better and more stable results in order to find input parameters settings that increase validity with respect to this particular scenario. This simple exploration demonstrated that ART could be used as an optimization tool to calibrate PAX by adjusting input parameter settings to closely meet subject matter experts' expected outcomes of the scenario in selected MoEs.

Conclusion

The work at IFW17 helped a lot to understand some of the difficulties or better challenges in the process of calibration. The team gained valuable insights into how to use Data Farming for the calibration of agent-based simulation models. The utilized methods of large scale calibration, micro scenario analysis and ART optimization proofed to be useful and supplemented each other very well.

Macro scenarios help to identify relevant scenario parameters as well as realistic ranges for these parameters that shall be used in terms of analyzing simulation results. Simulating micro scenarios can supplement effectively in understanding and narrowing in on plausible value ranges of parameters. The applicability of the results of the calibration performed in micro scenarios need to be checked by feeding them back into the respective large-scale scenario. Besides that, it could be demonstrated that PAX can be used for analyses integrating the optimization framework ART from DSO Singapore. In doing this, it was shown that ART can apply the implemented optimization functionalities outside of the originally intended *red teaming* purpose of the framework.

During the workshop, a preliminary approach for the process of calibrating an agent-based model like PAX was developed that of course needs to be further refined. But the first outcomes and findings are promising in a way, that it seems that even without the not-easy-to-get *hard data* – means empirically driven quantification of the relevant effects and outcomes of the model – a substantially increased quality of the model can be achieved by consequently calibrating the model using this workshop's approach including subject matter expertise and its related *qualitative data*.

EVALUATION OF ELECTRO-OPTICAL SENSOR SYSTEMS IN NETWORK CENTRIC OPERATIONS USING ABSEM 0.1 (HAYMANN, K. AND WAGNER, G., 2009).

The German Federal Office of Defense Technology and Procurement (BWB) is interested in analyzing the influence of networked sensors and effectors on military capabilities in network centric operations. Former studies showed, however, that the existing agent-based models are rather limited in terms of modeling and simulating complex technical systems on a sound physical basis.

For this reason, the BWB started developing a new agent-based model that aims to fulfill the requirements to be used for analyzing the combination of various sensor and effectors systems in NCO and taking into account underlying physical theories. A first prototype of this model (ABSEM version 0.1) was presented to the International Data Farming Community at IDFW17.

ABSEM Overview

The goal of ABSEM (Agent-based Sensor-Effector Model) is to design and establish a powerful agent-based tool for sensor-effector simulations on the technical and tactical level that can be used for high performance experimentation on the German Supercomputing / Data Farming cluster.

ABSEM concentrates on modeling the technical aspects in NCO, with the main focus on sensor and effector systems. For this reason, ABSEM integrates detailed physical theories when it comes to simulating the output of various sensors and when determining the effect of different weapon systems.

ABSEM version 0.1 mainly concentrates on modeling electro-optical sensors, including the “normal” human view, infrared devices and residual light amplifier. As input data the sensors evaluate information about the background, information about the target itself, but also atmospheric conditions. The sensors’ output is a list of perceived entities in the agent's field of view, along with the information if the target entity was detected, classified or even identified.

Objectives

The ABSEM implementation was only started in 2008, thus version 0.1 provides basic functionalities for modeling sensor and effector systems in NCO. The goal of the Data Farming experiments in the camp protection scenario was to scrutinize the ABSEM approaches implemented so far.

Overall, the team had the following goals:

- Review and face validate the first ABSEM prototype version 0.1.
- Conduct data farming experiments using common data farming tools.
- Identify needs for further work.
- Gain insight into other models (participation in plenary sessions).

Scenario

Using a camp protection scenario (Figure 7), the team's objective was to investigate the effect of different electro-optical sensor systems (human view, short wave, middle wave and long wave infrared) within a dynamic environment.

The IDFW17 scenario deals with the threat posed by micro-light aerial vehicles. The military camp is protected by several watch towers occupied with soldiers equipped with the electro-optical sensors mentioned above. The sentry reports any detected, classified or identified unit to the command center, which, in turn, decides how to proceed. Depending on the scenario setup (and the user-defined agent behavior), the own forces will fight the detected micro-light planes as soon as they were classified or wait for an identification (to reduce fratricide due to false interpretation).

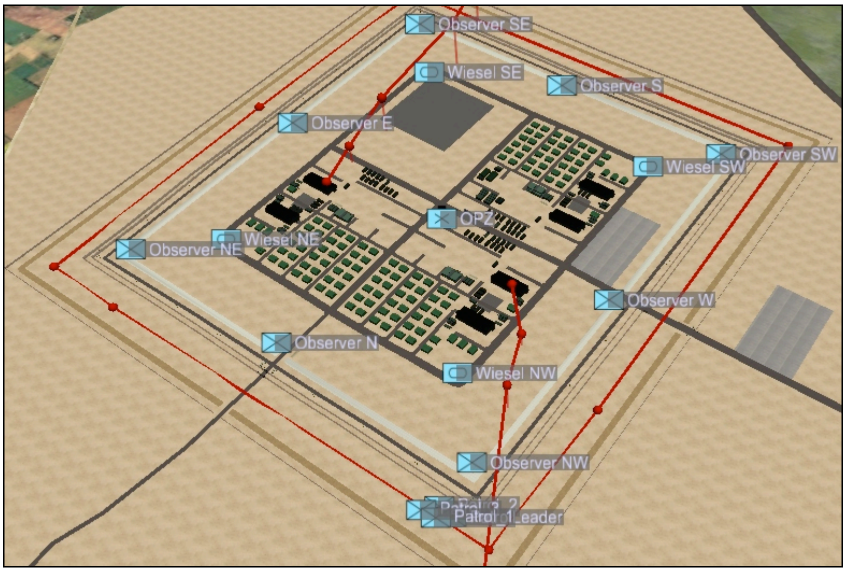

Figure 7: Scenario Camp Protection.

Team Activities

The team's main intention during IDFW17 was to investigate the effect of different electro-optical sensor systems. Therefore they examined different devices with either reflection- or temperature-based vision (see Table 4).

SensorID	Name	FieldOfViewH	FieldOfViewV
1	SC3000	45	45
2	Merlin	45	45
3	HumanEye_100Lux	45	45
4	HumanEye_10000Lux	45	45
5	Binoculars_100Lux	5	5
6	Binoculars_10000Lux	5	5
7	SWIRDevice_100Lux	45	45
8	SWIRDevice_10000Lux	45	45
9	HumanEye_1Lux	45	45
10	Binoculars_1Lux	5	5
11	SWIRDevice_1Lux	45	45

Table 4: Different types of sensors.

In several data farming experiments the different sensors' validity was checked. To ease things, the scenario described above was further simplified (see Figure 8).

In our scenario, the team examined the sensor's performance for different approach corridors for the hostile micro-light plane. By doing so, the team could analyse the effect of different terrain features.

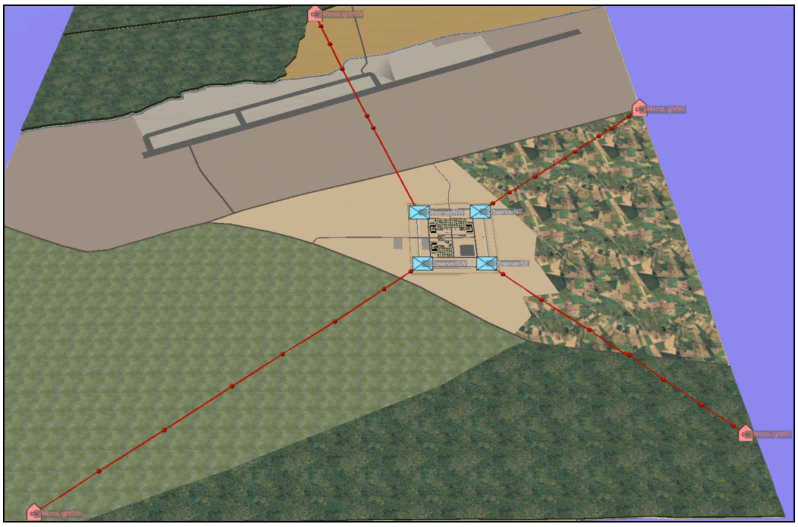

Figure 8: Simplified base case scenario.

Data Farming Experiments

After some preliminary experiments and looking at single simulation runs to get familiar with the model and its user interfaces, three different data farming experiments were performed.

In the first experiment we distinguished between day and night, thus varying the parameters for the luminance and the temperatures of the target object and the background.

In the second experiment the team had a look at different levels of attenuation. Thus representing the whole bandwidth from excellent to very bad weather conditions, that is from clear sky to dense fog.

Finally, the team was convinced themselves that the detailed physical ABSEM sensor delivers more realistic perception probabilities than a simple, purely probability based sensor. In a third experiment the team decided to equip the observer either with a detailed optical sensor or with a very simple and abstract sensor (comparable to the one implemented in MANA).

Data Farming Results

All three experiments were considering the distance at which the target object could be detected, classified or identified to measure the sensor's performance (plotted against the y-axis).

In the first experiment regarding the sensor performance during day and at night, we got the results we expected: the lower the luminance, the worse the perception (Figure 9).

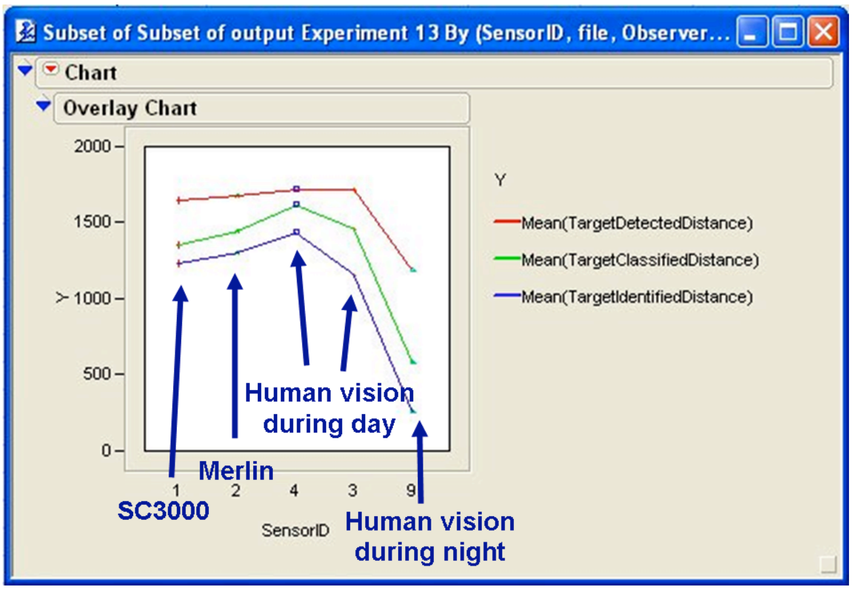

Figure 9: Experiment 1 — Distinction between day and night.

In the second experiment, the results were also quite obvious, which confirmed our sensor modeling approach: the higher the attenuation (that is the worse the weather conditions), the later the hostile entity was perceived (Figure 10).

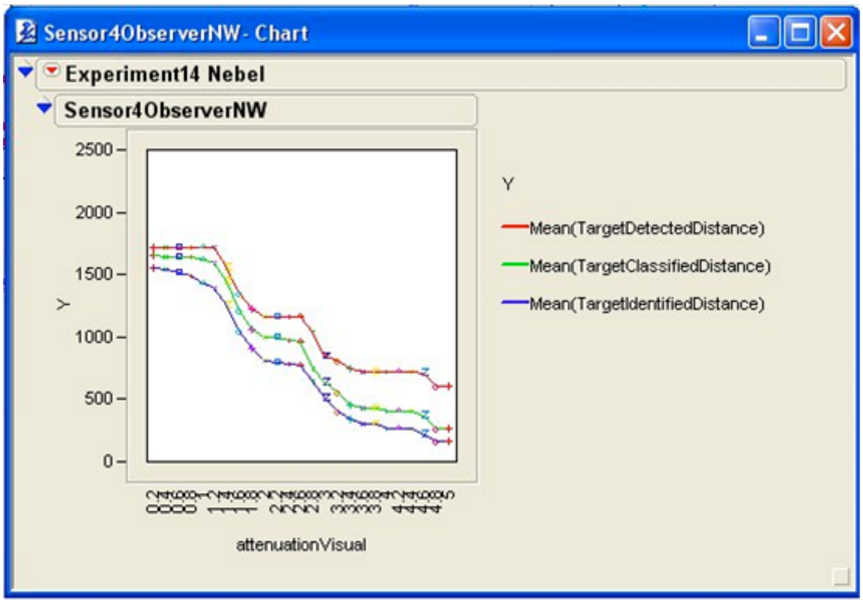

Figure 10: Experiment 2 — Different weather conditions.

Finally, with the third experiment, we simply wanted to demonstrate that the detailed physical modeling approach in fact results in much more realistic behavior. We compared a very simply modeled sensor with our detailed physical optical sensor. Since the simple sensor does not distinguish between different weather conditions or different times of the day, the detection distance was constant, of course (

Figure 11).

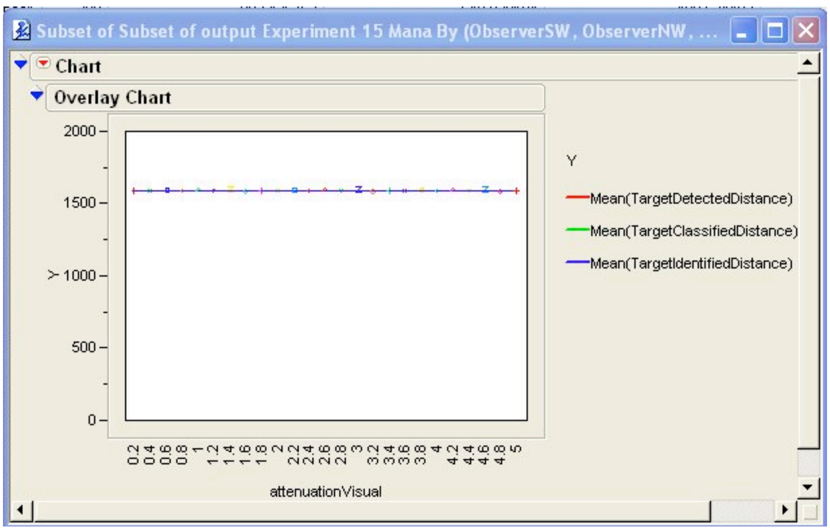

Figure 11: Experiment 3 — Simple and abstract sensor returns constant detection distance regardless of different weather conditions or times of the day.

Using the detailed physical optical sensors, by contrast, allows us to get much more reliable and realistic results. And that is essential when using ABSEM for analyzing the influence of complex technical systems (especially networked sensors and effectors) on military capabilities and the operational outcome.

Summary

The team's main intension was to review the first model prototype ABSEM v 0.1 (focussing on optical sensors). We could show that the physical sensor modelling approach we have chosen seems to be very promising. With the implemented optical sensors, the terrain features and atmospheric conditions are adequately considered.

Furthermore, we also succeeded in conducting data farming experiments using the Data Farming GUI. We were successfully running our experiments within a feasible amount of time on the German cluster (ABSEM currently runs significantly faster-than-real-time[1]). Overall we were able to confirm that we are on the right way!

INVITATION

This article has two purposes. The first is to describe the concept of Data Farming and give an overview of how it is being used worldwide. The second is to invite you to become part of our International Data Farming Community. We value openness, collaboration, and having fun in the process. By planting seeds of knowledge throughout the world we feel that we can grow the methods and tools to begin to provide answers to the difficult questions of our age.

We invite you to contact us, we invite you to use our tools and methods, and we invite you to join us in person at our next International Data Farming Workshop (IDFW 19) in Auckland, New Zealand from 08 through 13 November 2009 and/or IDFW 20 in Monterey, California in Spring 2010. There we will continue to strive to outline the landscapes of possibilities, discover surprises, and uncover those dynamic truths central to understanding questions that we share.

REFERENCES

Brandstein, A. and Horne, G., 1998. Asymmetric Evolutions and Revolutions in Military Analysis. In: Woodcock, Alexander and Davis Davis (eds.) The Cornwallis Group II: Analysis for and of the Resolution of Conflict. Cornwallis Park, Nova Scotia, Canada. The Lester B. Pearson Canadian International Peacekeeping Training Centre, The Canadian Peacekeeping Press.

Brandstein, A., 1999. *Operational Synthesis: Supporting the Maneuver Warrior*, briefing presented at the 2nd Annual Defense Planning & Analysis Society Symposium. Rosslyn, VA.

Haymann, K. and Wagner, G., 2009. *Scythe, Proceedings and Bulletin of the International Data Farming Community, Issue 5, Workshop 17*, The SEED Center for Data Farming, Monterey, CA.

Henscheid, Z., Middleton, D., and Bitinas, E., 2007. Pythagoras: An Agent-Based Simulation Environment, *Scythe Issue 1*: pp. 40-44. Monterey, California.

Hoffman, F. and Horne, G., 1998. *Maneuver Warfare Science 1998*. United States Marine Corps Project Albert. Quantico, Virginia.

[1] Execution time of 110000 runs on a 32-node cluster: 43 minutes, 15 seconds.

Horne, G., 1997. *Data Farming: A Meta-Technique for Research in the 21st Century*, briefing presented at the Naval War College. Newport, Rhode Island.

Horne, G., 1999. Maneuver Warfare Distillations: Essence Not Verisimilitude. *Proceedings of the 1999 Winter Simulation Conference*, eds. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans, pp. 1147-1151. Phoenix, Arizona.

Horne, G. and Leonardi, M., 2001. *Maneuver Warfare Science 2001*. United States Marine Corps Project Albert. Quantico, Virginia.

Horne, G. and Johnson, S., 2002. *Maneuver Warfare Science 2002*. United States Marine Corps Project Albert. Quantico, Virginia.

Horne, G. and Johnson, S., 2003. *Maneuver Warfare Science 2003*. United States Marine Corps Project Albert. Quantico, Virginia.

Horne, G. and Meyer, T., 2004. Data Farming, briefing presented at the *2004 INFORMS National Meeting*. Denver, Colorado.

Horne, G. and Meyer, T., 2004. Data Farming: Discovering Surprise. *Proceedings of the 2004 Winter Simulation Conference*, eds. R. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, pp. 171-180. Washington, D.C.

Horne, G. and Meyer, T. 2005. Data Farming Architecture. *Proceedings of the 2005 Winter Simulation Conference*, eds. M. E. Kuhl, N. M. Steiger, F.B. Armstrong, and J. A. Joines, 1082-1087. Orlando, Florida.

Kleijnen, J., Sanchez S., Lucas T., and Cioppa T., 2005. A User's Guide to the Brave New World of Designing Simulation Experiments, *INFORMS Journal on Computing*, 17(3): 263-289. Hanover, Maryland.

Lampe, T., Schwarz, G., and Wagner, G., 2007. PAX: Designed for Peace Support Operations, *Scythe Issue 2*: 43-48. Monterey, California.

Lauren, M., McIntosh, G., Galligan, D., Anderson, M., 2007. Recent Developments in the MANA (Map Aware Non-uniform Automata) Agent-based Model, *Scythe Issue 1*: 38-39. Monterey, California.

Lawler, M., 2005. Data Farming Cultivates New Insights. *SIGNAL, International Journal of the Armed Forces Communications and Electronics Association, June 2005*. Fairfax, Virginia.

McDonald. M., 2008. *The Use of Agent-Based Modeling and Data Farming for Planning System of Systems Tests in Joint Environments*, briefing presented at the 76th Military Operations Research Society Symposium. New London, Connecticut.

Meyer, C. and Davis, S., 2003. *It's Alive*. Ernst and Young Center for Business Innovation, Cambridge, Massachusetts.

Meyer, T. and Horne, G., 2007. *Scythe, Proceedings and Bulletin of the International Data Farming Community, Issue 1, Workshop 13*, The SEED Center for Data Farming, Monterey, California.

Meyer, T. and Horne, G., 2007. *Scythe, Proceedings and Bulletin of the International Data Farming Community, Issue 2, Workshop 14*, The SEED Center for Data Farming, Monterey, California.

Meyer, T. and Horne, G., 2008. *Scythe, Proceedings and Bulletin of the International Data Farming Community, Issue 3, Workshop 15*, The SEED Center for Data Farming, Monterey, California.

Meyer, T. and Horne, G., 2008. *Scythe, Proceedings and Bulletin of the International Data Farming Community, Issue 4, Workshop 16*, The SEED Center for Data Farming, Monterey, California.

Meyer, T. and Horne, G., 2008. *Scythe, Proceedings and Bulletin of the International Data Farming Community, Issue 5, Workshop 17*, The SEED Center for Data Farming, Monterey, California.

Sanchez, S. and Lucas, T. 2002. Tutorial: Exploring the World of Agent-based Simulations: Simple models, Complex Analyses. *Proceedings of the 2002 Winter Simulation Conference*, eds.: E. Yucesan, C. H. Chen, J. L. Snowdon, J. M. Charnes, pp. 116-126. San Diego, California.

Schwierz, K., 2008. *Chronicle of Data Farming: The Story of Success*, briefing presented at IDFW 16. Monterey, California.

Seichter, S. and Schwarz, G., 2009. *Scythe, Proceedings and Bulletin of the International Data Farming Community, Issue 5, Workshop 17*, The SEED Center for Data Farming, Monterey, California.

Upton, S., 2004. *User's Guide: OldMcData, the Data Farmer, Version 1.0*. United States Marine Corps Project Albert. Quantico, Virginia.