Proceedings of the 2003 Winter Simulation Conference
S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice, eds.

FREQUENCY-BASED DESIGNS FOR TERMINATING SIMULATIONS: A PEACE-ENFORCEMENT EXAMPLE

Susan M. Sanchez

Operations Research Department and
Graduate School of Business & Public Policy
Naval Postgraduate School
Monterey, CA 93943-5219, U.S.A.

Hsin-Fu Wu

PCU JIMMY CARTER (SSN 23)
United States Navy
Groton, CT 06340-4909, U.S.A.

ABSTRACT

In recent years, the U.S. Marine Corps has begun developing an infrastructure for applying agent-based models and simulation, computing power, and data analysis and visualization technologies to help answer complex questions in military operations. Factor screening approaches are of particular interest, since even relatively simple agent-based models may have hundreds (or even thousands) of inputs that can be varied. We describe a new experimental design, called a frequency-based design, that can be used for exploring the behavior of terminating simulations. We apply this to a model of a peace-enforcement operation. We examine the behavior of four performance measures (including two attrition ratios) and discuss how the results confirm and complement earlier findings. We conclude with a brief discussion of issues that merit further investigation.

1 INTRODUCTION

Our motivation for exploring this approach arose out of work we have been doing for the U.S. Marine Corps under their "Project Albert" umbrella. This multi-year, multi-national effort attempts to exploit advances in three core disciplines: 1) Agent-based models and simulations; 2) Computing power; and 3) Data visualization. *Data farming* (Brandstein and Horne, 1998) is the application of these disciplines to help answer complex questions in military operations. The four principal processes of data farming are fertilization, cultivation, planting, and harvesting. *Fertilization* means providing military professionals and other experts with ideas on how to capture important aspects of conflict that have not been taken into account in the past, such as morale, leadership, timing, intuition, adaptability, etc. *Cultivation* means receiving ideas from these professionals about what might be important in a given situation. *Planting* means incorporating these ideas into models, to the extent possible, and running the models over a wide variety of possibilities by varying simulation parameters (also called factors). Finally, innovative techniques for understanding scientific data are used when *harvesting* the model output. Just as the farmer grows crops to meet the needs of the consumers, who are hungry for food, the data farmer grows data to meet the needs of the ultimate decision-makers, who are hungry for answers.

The propensity to produce large multi-dimensional data sets with several measures of performance (MOPs) is inherent in data farming. However, care must be taken when generating data, because the time required to examine all potential factor level combinations grows exponentially with the number of factors investigated. Furthermore, because Project Albert uses agent-based simulations to model military operations, some of the factors are only notional representations of human thought or behavior—such as morale, unit discipline, leadership, or aggressiveness. The data farming environment is thus best viewed as one that may provide the decision-makers with qualitative insights rather than numerical predictions (Lucas et al., 2002, 2003).

The vagueness associated with interpreting the factors can complicate the task of capturing and communicating the essence of the data set. However, gaining insight into the model's behavior is challenging for any simulation with a large number of factors, particularly if interaction effects among two or more of these factors are possible. Brute force methods are not practical unless the number of factors is small. Trial-and-error methods are unlikely to provide the analyst with an understanding of how the MOPs are affected by the input factor settings. Instead, systematic experimental designs are needed in order to efficiently generate data that can be used to provide insights into the model's behavior and guide further investigations.

The type of design that is most appropriate depends on both the number of factors and the nature of the response surface. For an overview of the possibilities, we refer the reader to Sanchez and Lucas (2002) or Kleijnen et al. (2003). Clearly, no single design is best for all situations. However, the ability to use prior information can dramatically reduce the need for additional runs. This means that sequential designs have distinct advantages over non-sequential designs.

In this paper, we propose a frequency-based design (FBD) that is appropriate for analyzing terminating simulations. We present some background material on spectral analysis in Section 2, and describe FDB in Section 3. An agent-based implementation of a peace-enforcement scenario is explored using FBD in Section 4. We conclude with a brief discussion of some issues that merit further investigation.

2 BACKGROUND

In this section we present a very brief overview of some important concepts in spectral analysis. We also summarize a technique called frequency domain experimentation, which was developed by Schruben and Cogliano (1987) as a factor-screening approach for non-terminating simulations.

2.1 Spectral Analysis

Spectral analysis has long been used for identifying the cyclic components of time-series data (Chatfield, 1996).

Let $Y = [y_1\ y_2\ \ldots\ y_N]^{\mathrm{T}}$ be an indexed set of observations. For notational convenience, assume N is even and let $H = N/2$. We can decompose Y into its cyclic components by frequency as $Y = \mathrm{A}\theta$, where $\theta = \left[\mu\ \alpha_1\ \beta_1\ \ldots\ \alpha_H\ \beta_H\ \alpha_{N/2}\right]^{\mathrm{T}}$ and $\mathbf{A}$ is the matrix

$$\mathrm{A} = \begin{pmatrix} 1 & \ldots & \cos(\omega_i) & \sin(\omega_i) & \ldots & \cos(\pi) \\ 1 & \ldots & \cos(2\omega_i) & \sin(2\omega_i) & \ldots & \cos(2\pi) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \ldots & \cos(N\omega_i) & \sin(N\omega_i) & \ldots & \cos(N\pi) \end{pmatrix}. \quad (1)$$

The pairs of cosine and sine terms correspond to $\omega_i = 2\pi i/N$ for $i = 1, \ldots, H$. These frequencies are expressed in radians per observation, e.g., an oscillation of one cycle per observation equals 2π radians per observation. Furthermore, for any discretely-sampled signal it is sufficient to display only frequencies ranging from $[0, \pi]$ in the spectrum because the highest observable frequency is π radians per cycle (equivalently, one-half cycle per observation).

The Fourier transform of Y is $\hat{\theta} = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}Y$. Moving from the Fourier transform to the (Fourier) spectrum involves squaring the estimated coefficients for the sine and cosine terms, and summing them by frequency. The spectrum components are thus $(\alpha_i^2 + \beta_i^2)$ for $i = 1, \ldots, H$ and α_i^2 for $i = N/2$. It can be demonstrated that

$$\frac{1}{N}\sum_{i=1}^{N}(y_i - \overline{y})^2 = \sum_{i=1}^{N/2}\left(\alpha_i^2 + \beta_i^2\right) + \alpha_{N/2}^2, \qquad (2)$$

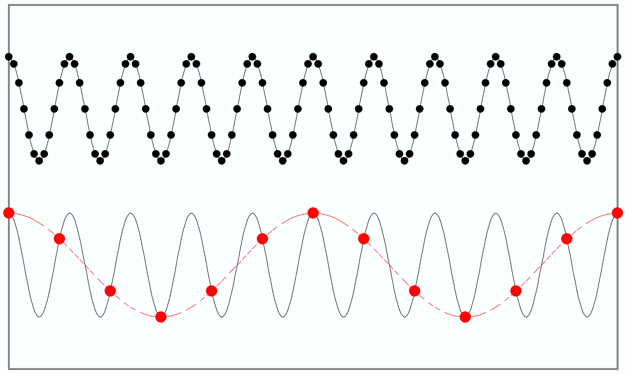

Figure 1: Example of Aliasing

i.e., the Fourier spectrum partitions the variance. Under mild assumptions (Chatfield, 1996), the estimated spectral coefficients have a chi-squared distribution.

While the regression representation makes it clear how the partitioning works, other computationally efficient methods can be used to estimate the spectrum. The Wiener-Khintchine theorem relates the Fourier spectrum of the model to the Fourier transform of the autocovariance function of the observations in the data set. Autocovariance is a measure of the covariance of a sequence of observations with each other. For stationary processes, autocovariance generally diminishes as the observations become sufficiently far apart. A technique called windowing uses only a specified number of observations (M, called the window size) to estimate the spectrum efficiently. A common choice is to select M to be proportional to $\sqrt{N}$, although other values of M are possible so long as that the ratio $M/N \to 0$ as $M,\ N \to \infty$. A more thorough explanation of Fourier analysis can be found in Chapter 7 of Chatfield (1996).

One other important issue is frequency *aliasing*. If a frequency oscillates at a rate higher than one-half cycle per observation, it appears to be "folded" back into the range $[0, \pi]$ when sampled discretely. Figure 1 illustrates this behavior. Both signals completes 10 full cycles in the graph. The sampling rate in the upper figure is high enough (1/12 cycle per observation) that an observer would view the true frequency. However, aliasing occurs in the lower graph where the sampling rate is 5/3 cycles per observation. The sampled data appear to complete only two cycles, for an apparent frequency of 1/3 cycle per observation.

2.2 Frequency Domain Experiments for Non-Terminating Simulations

The idea of oscillating input factor levels was first explored by Schruben and Cogliano (1987). Their approach, called frequency domain experimentation (FDE), investigates the impact of several input factors on system performance by varying input factor levels within the course of a single, very long run called the *signal run*. All input factors are held constant at nominal levels during another run of the same length, called the *noise run*. After truncating both the signal and the noise runs to remove the initial transients, the Fourier spectra of the two output streams are obtained. The ratios of the signal spectrum to the noise spectrum are computed by frequency. High values or "spikes" in these signal-to-noise ratios (SNRs) indicate that the associated input term is an important contributor to the output behavior.

The frequency at which a factor is varied is called its *driving frequency*. In the frequency domain, the spectrum displays all frequencies contributing to the variations in the response. Furthermore, the indicator frequencies for higher-order effects of the oscillated factors on response show up at well-defined locations. For example, suppose factor X_1 is assigned a driving frequency ω_1. The main effect of X_1 on the response spectrum has an indicator frequency of ω_1. If factor X_1 has a quadratic effect on the response, this is associated with an indicator frequency of $2\omega_1$. Similarly, the nth-order effect of X_1 on the response has an indicator frequency of $n\omega_1$ (or its alias) in the response spectrum. For interaction terms, the indicator frequencies are the sums and differences of the driving frequencies. For example, suppose there is a second-order interaction effect on the response from two factors, X_1 and X_2, where X_1 and X_2 are assigned driving frequencies of ω_1 and ω_2, respectively. The second-order interaction term in the response has two indicator frequencies: $\omega_1+\omega_2$ and $\omega_1-\omega_2$, respectively. If $\omega_1+\omega_2$ falls outside the interval $[0, 2\pi]$, then the associated indicator frequency is the alias (and similarly for $\omega_1-\omega_2$).

Sanchez and Buss (1987) provide a mathematical model formulation which explains how and why FDEs work. Jacobson, Buss and Schruben (1991) provide an algorithm which assigns driving frequencies to the factors to allow investigation of full second-order or third-order models. Sanchez and Konana (2000) investigate how the total data collection effort should be allocated between the signal and noise runs, with the goal of improving the efficiency of FDEs.

3 FREQUENCY-BASED DESIGNS FOR TERMINATING SIMULATIONS

We now describe a new data-farming approach: the use of frequency-based designs (FBDs) for terminating simulations. While we make use of many of the building blocks of FDE, the FDE approach cannot be directly applied to terminating simulations. For some simulation models, such as the peace-enforcement application of Section 4, long runs are not possible. For others, such as rare-event simulation, the time until termination may itself be the performance measure.

Our frequency-based designs are similar to the signal runs for FDEs in terms of factor level selection, etc. They differ in several notable ways. First, the input factor levels are held constant for the duration of each run, and we obtain only one data summary from each run that will be used for analysis purposes. Second, all runs are independently seeded—which means there are no serially correlated errors in the run-to-run results. Third, since there are no concerns about initial bias, we can use all of the data generated by the simulation. Fourth, our implementation exploits the power of supercomputers by parallelizing the data farming process. While this last point is not a requirement for using FBDs, it does have practical benefits by reducing the total length of time necessary to acquire the data.

We now provide enough detail for the interested reader to generate and apply FBDs to terminating simulations. Alternatively, software programs to perform the design and analysis are available (Sanchez, 2002). Our example assumes we are interested in fitting a second-order model, i.e., one where quadratic and/or two-way interactions may be present. A similar approach can be used for higher-order models, although the required number of runs increases.

Suppose there are k input factors $X_1, \ldots, X_k$ we wish to investigate. To specify the design we must determine the number of runs (N) and the factor levels at each run ($X_{i,t},\ i = 1, \ldots, k;\ t = 1, \ldots, N$). Our procedure follows.

1. Specify the input factors $X_1, \ldots, X_k$, along with middle, minimum and maximum values of interest for the experiment ($m_i,\ m_i - a_i, m_i + a_i$, respectively).
2. Determine a set of driving frequencies $\omega_1, \ldots, \omega_k$ that will allow identification of any important terms in a 2nd order metamodel:
$$\hat{Y}_t = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i X_{i,t} + \sum_{i=1}^{k} \sum_{j=i}^{k} \hat{\beta}_{ij,t} X_{i,t} X_{j,t}.$$
Let f denote the lowest common denominator of the frequencies when expressed in cycles per observation, and let $p = 2k+[k(k-1)/2]$ represent the total number of terms in the full model.
3. Set $N = cf$ for some number of replications c. Then set
$$X_{i,t} = m_i + a_i \cos(\omega_i t), \qquad t = 1, \ldots, N.$$
for $i = 1, \ldots, k$ and $t = 1, \ldots, N$. Y_t is the output of run t.
4. Take the Fourier transform of the output Y to obtain $\hat{\theta}$. Under mild assumptions, the components of the spectrum will have chi-squared distributions (let ν denote the degrees of freedom). The variance in Y is now partitioned into components for the frequencies. Let S_ω denote the spectrum evaluated at frequency ω (or its alias in the range $[0, \pi]$.

5. Sum the spectral values at all non-indicator terms to obtain an estimate of the noise component. Call this value V. The number of df associated with V is $r = \nu\,[0.5(f-1)-p]$.
6. Compute signal-to-noise ratios (SNRs) by mapping the indicator frequencies to the metamodel terms and dividing by the noise component.
   - For the main effect of X_i,
   $$\mathrm{SNR} = \frac{S_{\omega_i}/\nu}{V/r}.$$
   - For the quadratic effect X_i^2,
   $$\mathrm{SNR} = \frac{S_{2\omega_i}/\nu}{V/r}.$$
   - For the interaction effect of X_i and X_j,
   $$\mathrm{SNR} = \frac{(S_{\omega_i+\omega_j} + S_{\omega_i-\omega_j})/2\nu}{V/r}.$$
7. Graph or tabulate the SNRs.
8. For an overall test at level α, compare the SNRs for the main and quadratic effects to the critical value $F^{(1-\alpha/p)}_{(\nu,r)}$. and compare the SNRs for the interaction effects to the critical value $F^{(1-\alpha/p)}_{(2\nu,r)}$. Alternatively, particularly if many SNRs are statistically significant, a qualitative assessment identifies those terms with the largest SNRs as the most important.
9. If no terms are statistically significant, either stop (and conclude that the selected input factors do not affect the response over the ranges examined) or increase c and go back to Step 3 to collect more data.

Software useful for designing and analyzing FBDs (or FDEs) is available in the Java© programming language (Sanchez, 2002). The `Design` program requires the number of factors as the input, and returns a set of frequency assignments as in Table 1. The `Fourier` program requires the following inputs: the number of frequencies into which the response is to be partitioned, the window size, the type of windowing, and the number of observations in the input data set. The program then estimates the spectrum of the observations and produces a response spectrum. The program automatically adds one more partition for the zero frequency, that corresponds to the constant term in the regression model. Thus, the spectral power at the zero frequency signifies the contribution of the constant term in the regression model to the response.

Note that the experimental units are independent, by construction, since we use different random number seeds for each experiment, It follows from the Wiener-Khintchine theorem (Chatfield, 1996) that the spectrum of independent observations is flat. Thus, under the null hypothesis that there are no factor effects, the heights of the indicator and non-indicator frequencies have the same expected value for all ω, and so the expected value of each SNR is equal to 1.

4 A PEACE-ENFORCEMENT EXAMPLE

Peace enforcement is a critical component of current and future military operations. According to the U.S. Army Field Manual 100-23 (Department of the Army, 1994), peace enforcement is “the application of military force or the threat of its use, normally pursuant to international authorization, to compel compliance with generally accepted resolutions or sanctions. The purpose of peace enforcement is to maintain or restore peace and support diplomatic efforts to reach a long-term political settlement."

There are many ways in which peace-enforcement operations can be investigated. We consider a scenario developed by Cioppa (2002), who developed and used nearly-orthogonal Latin hypercubes to explore the model’s performance. The scenario was deemed doctrinally correct and plausible by the U.S. Army Infantry Simulation Center at Fort Benning, Georgia. The scenario was implemented in MANA (Map Aware Non-uniform Automata)—an agent-based modeling platform that was developed for the New Zealand Army and Defence Force (Stephen and Lauren, 2001). It has a graphical user-interface for specifying initial conditions and trigger states of the agents, as well as for animating the simulation run. MANA also offers the user the ability to specify levels of input parameters easily from a formatted input file. MANA is one of Project Albert’s data-farmable suite of modeling platforms.

4.1 Scenario

Figure 2 illustrates the initial positions of the agents for a single run of the software. The ellipses indicate four areas of operations (AOs). Clockwise from the center left are AO Rattler, AO Python, AO Cobra and AO Boa. Blue’s mission is to clear AO Cobra within the next two hours in order to facilitate United Nations (UN) food distribution and military convoy operations. Blue uses a light infantry platoon composed of three nine-man rifle squads and a platoon headquarters (HQ) of seven soldiers containing two machine guns. Their movement scheme is one squad up and two squads back, with the platoon HQ following the lead squad (squad 2). The lead squad’s task is to conduct a movement to contact with the purpose of clearing AO Cobra. Their follow-on task is to clear AO Cobra for subsequent UN food distribution and military operations. Squad 1’s task is to follow and support the lead squad with the purpose of clearing AO Cobra. Their follow-on task

Table 1: Input Factors for the Peace-Enforcement Scenario

Factor Name	Description	Range	Frequency Assignments Batch 1	Batch 2	Batch 3
U	Blue Squad 1 Contact Cohesiveness: controls the propensity to remain with the squad when it encounters Red agents.	−64, +64	1/81	29/81	10/81
F	Blue Lead Squad Contact Cohesiveness: controls the propensity to remain with the squad when it encounters Red agents.	−64, +64	4/81	1/81	17/81
G	Blue Squad 3 Injured Cohesiveness controls the propensity to remain with the squad when one or more members are injured.	−64, +64	10/81	4/81	29/81
P	Movement speed for all Blue agents.	72, 200	17/81	10/81	1/81
V	Red Aggression: controls Red's propensity to pursue a perceived threat.	−64, +64	29/81	17/81	4/81

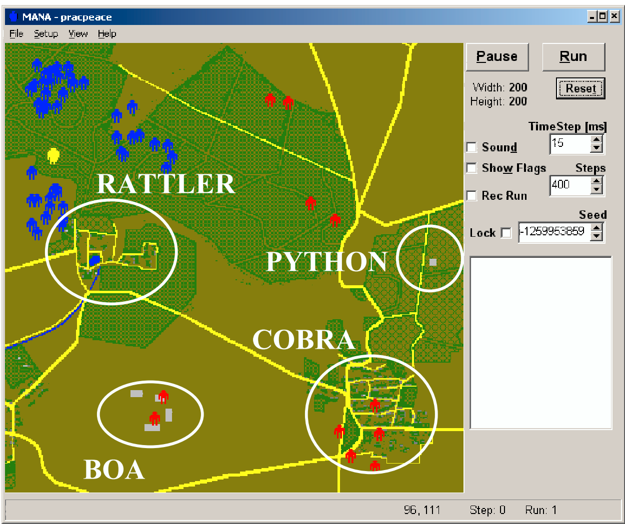

Figure 2: Initial Graphical Depiction of MANA Peace-enforcement Scenario

is to clear AO Python. Squad 3's task is to follow and support the lead squad with the purpose of clearing AO Cobra. Their follow-on task is to clear AO Boa (a small urban area with four building structures). After the lead squad clears AO Cobra, the platoon HQ moves to AO Boa to provide supporting fires.

Red has a five-member element located near AO Cobra and two two-member elements patrolling along the movement routes of the two Blue supporting squads. Red also has a two-member element near AO Boa. A Yellow three-member element begins in the midst of the Blue forces in the upper left of Figure 2. Yellow is initially non-hostile, but becomes hostile after discovering that there is no potable water in the vicinity of AO Rattler. Yellow then seeks small arms from the vicinity of AO Boa and moves to the vicinity of AO Python.

The devised scenario runs for a user-specified time interval before terminating. The scenario is challenging since the Blue force is subjected to a series of encounters with the Red force and an originally non-hostile force (Yellow) turns hostile as the scenario progresses. In this scenario, Red is aggressive and exchanges fire with Blue rather than running away. Modifying the agent personalities, e.g., making some of the Red agents less likely to approach large groups of Blue soldiers, could lead to another scenario with quite different behavior. The ability to make such changes quickly is one of the benefits of an agent-based modeling platform such as MANA.

4.2 Input Factors

Because our primary interest is to determine the feasibility of applying FBD in a data-farming environment, rather than conduct a thorough exploration of a particular scenario, we vary only five factors. These are the factors Cioppa (2002) found to be most influential when he used a nearly orthogonal Latin hypercube design to examine 22 factors. We leave the remaining factors at their nominal values (i.e., the base settings) in all runs of the scenario. Table 1 lists the factors, along with brief descriptions and the ranges over which they are varied.

4.3 Output Responses

At the time this study was conducted, MANA's outputs were limited to the numbers of agents "killed" during the user-specified number of time steps before the simulation terminates. If Blue is using non-lethal weapons, then the number of Red "killed" corresponds to those incapacitated for the remainder of the operation. Wagner, Sanders, and

Mylander (1999) provide guidance on choosing an appropriate measure of effectiveness (MOE): it must be quantitative, measurable, reflect both the benefits and penalties of a particular course of action, and a significant increase (decrease) must correspond to a significant improvement (worsening) in achieving the decision-maker's objective. We choose not to limit our investigation to a single MOE. Let N_B and N_R denote the numbers of Blue and Red agents in the scenario. Our four performance measures follow.

1. K_B: the number of Blue killed,
2. K_R: the number of Red killed,
3. $ER = K_R/(K_B+1)$: the modified exchange ratio,
4. $FER = K_B/K_R$: the fractional exchange ratio.

K_B and K_R are the MANA outputs, while ER and FER are computed from these output streams. The denominator in ER is K_B+1 instead of K_B because, in some of the runs, no Blue agents were killed. A similar adjustment could be made for FER, but was unnecessary for our scenario since Red always sustained losses.

4.4 FBD Implementation

We want the ability to fit a second-order polynomial metamodel to the output responses. As indicated by the `Design` program (Sanchez, 2002), the spectrum must be partitioned into 81 discrete frequencies to accommodate unique indicator frequencies for each of the 5 main effects, 5 quadratic terms, and 20 effects corresponding to the 10 interaction terms. The three sets of frequency assignments are provided in the last three columns in Table 1. We consider each set to be one batch of five hundred "rows" of eighty-one "genetically engineered strains" that we plant using the MANA distillation. In other words, the frequency assignments remain the same for all factors within each planted batch of data. The settings at the beginning of each oscillation are assigned to their respective maxima as in equation (3). We replicate the set of eighty-one experimental units five hundred times in each batch since we know (from Cioppa, 2002) the results are highly variable. We also plant three batches of data in the data landscape since we have three frequency assignment schemes for the factors. The total number of experimental units grown in our FBD is 121,500. It took about 31 hours to complete the runs.

4.5 Results

We harvest the output data sets and process the two MOPs and the two attrition ratios for all three batches through the spectral analysis program, `Fourier`. We then use the list of indicator frequency mappings (provided by `Design` along with the driving frequencies) to collapse the resulting response spectrum by the corresponding terms in the regression model.

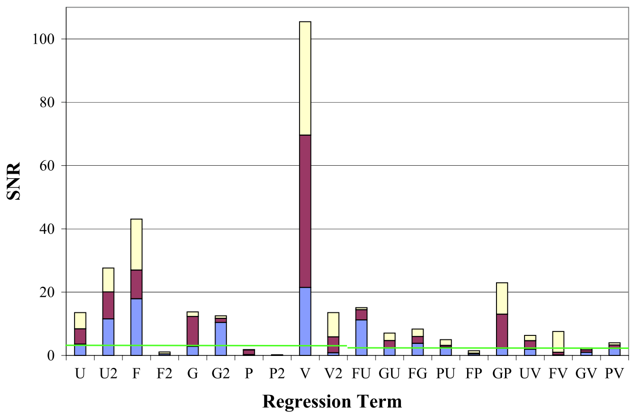

Figure 3: Signal-to-Noise Ratio for Blue Killed

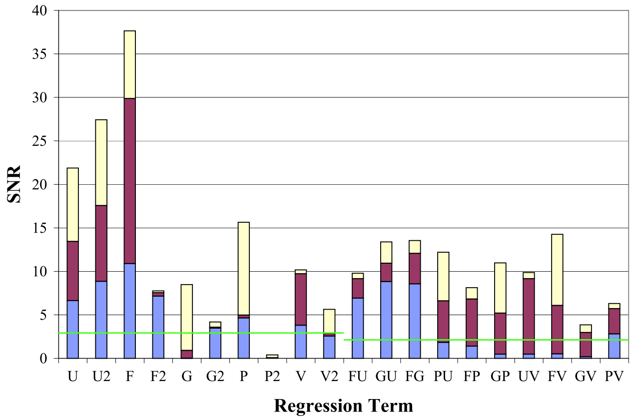

Figure 4: Signal-to-Noise Ratio for Red Killed

As Step 6 in our procedure describes, we then map the indicator frequencies in each batch to the associated terms in our regression model. Because the spectrum is a partition of discrete frequency bins, we present the spectra as stacked bar graphs rather than continuous linear graphs. The combined SNR graphs in Figures 3-5 illustrate the results for K_B, K_R, and ER, respectively. The SNR graph for FER (not shown) is quite similar to that for ER. The horizontal lines in these figures indicate the critical F-values for testing the statistical significance of the terms. Recall that the interaction terms have twice the degrees of freedom as do the main and quadratic effects since there are two indicator frequencies for each interaction term.

The three shades in each term correspond to contributions from the three different batches. Since the batches are independent, we could have chosen to run a single batch. (This differs from the FDE approach, where the system may have inherent tendencies to dampen or magnify effects at specific frequencies.) However, running three batches is informative. The differences in spike sizes across the three batches for several terms either indicates that the system is highly variable, or else highlights the fact that certain sam-

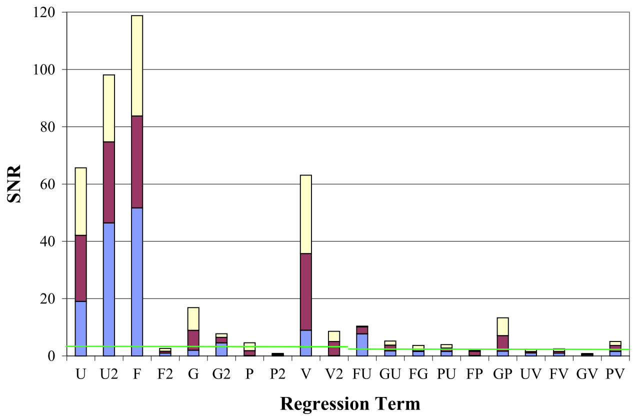

Figure 5: Signal-to-Noise Ratio for Modified Exchange Ratio (Red/Blue + 1)

pling patterns (frequencies) are more powerful than others for detecting specific types of effects. Were we to perform secondary experiments, we would continue to use a total of at least 121,500 observations rather than relying on results from a single batch of 40,500 new observations. Recall that the 5 factors we vary are those that Cioppa (2002) found to have the strongest effects on the (unmodified) exchange ratio K_R/K_B after using a nearly-orthogonal Latin hypercube design to plant the data. His regression model includes main effects for all five factors (F, G, P, U, and V), a quadratic term for U, and interaction terms for GU and FV. (It also includes four main effects and interactions involving factors we did not vary.) Figure 5 illustrates that the five main effects, one quadratic effect, and two interactions identified by Cioppa were also statistically significant in our FBD. We also found that G and V have quadratic effects on *ER*, and that other interaction terms are present. It is not surprising that our results are somewhat different. We varied fewer factors, modified the performance measure slightly, and used different ranges (in part to avoid a range of movement speeds where he found that the simulation model was broken).

Our results also highlight the importance of examining several different performance measures if the choice of MOE is unclear. Factor V (Red aggression) is by far the dominant factor in the SNR for K_B (Figure 3). However, the SNR for K_R is quite different (Figure 4), with V having the second-lowest main effect among the five factors. The SNR for K_R also indicates that a number of interaction terms involving F, G, P and U contribute as much or more to the variability in K_R as does factor V. The SNRs for *ER* and *FER* show F has the largest main effect, U and V have similar main effects, and U has a very strong quadratic effect as well. Other terms are statistically significant, but these four account for the vast majority of the variability in Y. So, an analyst using K_B as the performance measure might conclude that since Red aggression is the primary driver, there is little that can be done in planning the operation to reduce the likelihood of Blue losses. An analyst focusing on Red losses would see these are affected by Blue's contact cohesiveness (factors U and F) and Blue's movement speed (factor P). In both cases, the number killed for one affiliation is a function of the other side's actions, not their own. However, when looking at the attrition ratios, characteristics of squads from both sides contribute to the outcome.

Since the spectral terms are proportional to sums of *squared* metamodel coefficients, further analysis would be needed to determine appropriate levels for maximizing or minimizing the various performance measures. This can be accomplished using multiple regression without generating additional data. For numerical stability, it is best to use standardized independent variables X_i^* where

$$X_i^* = \alpha_i \cos(\omega_i t) \qquad t = 1, \ldots, N$$

(or $X_i^* = \cos(\omega_i t)$) for $i = 1, \ldots, k$. Quadratic and interaction explanatory variables can be obtained by multiplying the appropriate X_i^*.

5 CONCLUDING REMARKS

We have presented a frequency-based design (FBD) approach that can be used to identify important determinants of the performance of terminating simulations. Our work was motivated by the need to develop efficient, effective tools for planting data in a data-farming environment. As a proof of concept for the potential utility of this approach, we illustrated FBDs using an agent-based model of a peace-enforcement operation. Our results were in consonance with those obtained in an earlier (and broader) investigation that used the exchange ratio as the performance measure. We also considered three other performance measures, and found that looking at multiple performance measures gave additional insights into the relative importance of the factors. Therefore, we concluded that FBD is not only a feasible method for data farming, but also a useful technique for factor screening that is easy to generate. We are currently examining several issues regarding FBDs in order to further increase their utility. First, we are looking at sequential analysis and display of the results. This may be important either when the individual simulation runs take more time, or when we the number of factors is large. Information obtained early on the experiment may allow the analyst greater flexibility in changing the driving frequencies and/or factor ranges as initial results indicate either very strong or very weak effects. Second, Wu (2002) considered ways of sonifying the output. While further work is needed, sonification may be beneficial because it is another channel of information that could augment visual displays. Finally, we are assessing the use of expert opinion about what factors are likely to be the most important in assigning driving frequencies. Our goals are to see whether this additional information can be exploited to reduce the data requirements.

ACKNOWLEDGMENTS

This work was supported in part by grants from the United States Marine Corps Combat Development Command and the United States Marine Corps Warfighting Laboratory.

REFERENCES

Brandstein, A. G. and G. E. Horne. 1998. Data farming: a meta-technique for research in the 21st century. In *Maneuver Warfare Science*, eds. F. G. Hoffman and G. E. Horne. Quantico, Virginia: United States Marine Corps Combat Development Command.

Chatfield, C. 1996. *The Analysis of Time Series: An Introduction*. 5th ed. London: Chapman & Hall.

Cioppa, T. M. 2002. Efficient nearly orthogonal and space-filling experimental designs for high-dimensional complex models. Doctoral dissertation, Operations Research Department, Naval Postgraduate School, Monterey, California. Available online via <http://theses.nps.navy.mil/02sep_Cioppa_PhD.pdf> [accessed May 21, 2003].

Department of the Army. 1994. *Field Manual 100-23 Peace Operations.* Washington, D.C: Department of the Army.

Jacobson, S., A. H. Buss, and L. W. Scruben. 1991. Driving frequency selection for frequency domain simulation experiments. *Operations Research* 39(6), 917–924.

Kleijnen, J. P. C., S. M. Sanchez, T. W. Lucas, and T. M. Cioppa. 2003. A user's guide to the brave new world of simulation experiments. Working paper, Tilburg University, Tilburg, The Netherlands.

Lucas, T. W., S. M. Sanchez, L. P. Brown, and W. C. Vinyard. 2002. Better designs for high-dimensional explorations of distillations. In *Maneuver Warfare Science 2002*, eds. G. Horne and S. Johnson, 17–46. Quantico, Virginia: USMC Project Albert.

Lucas, T. W., S. M. Sanchez, T. M. Cioppa, and A. I. Ipekci. 2003. Generating hypotheses on fighting the global war on terrorism. In *Maneuver Warfare Science 2003*, eds. G. Horne and S. Johnson, Quantico, Virginia: USMC Project Albert, forthcoming.

Sanchez, P. J. 2002. *Design.java* and *Fourier.java*. Pacific Grove, California: Indalo Software.

Sanchez, P. J. and A. H. Buss. 1987. A model for frequency domain experiments. In *Proceedings of the 1987 Winter Simulation Conference*, eds. A. Thesen, H. Grant, W. D. Kelton, 424–427. Piscataway, New Jersey: Institute of Electrical and Electronic Engineers.

Sanchez, S. M. and T. W. Lucas. 2002. Exploring the world of agent-based simulations: simple models, complex analyses. In *Proceedings of the 2002 Winter Simulation Conference*, eds. J. L. Snowdon, J. Charnes, C.-H. Chen, and E. Yücesan, 116–126. Piscataway, New Jersey: Institute of Electrical and Electronic Engineers.

Schruben, L. W. and V. J. Cogliano. 1987. An experimental procedure for simulation response surface metamodel identification. *Communications of the ACM*, **30**(8), 716–730.

Stephen, R.T. and M.K. Lauren. 2001. *MANA Map Aware Non-Uniform Automata Version 1.0 User's Manual*. Land Operations Division, Defence Science and Technology Organisation, Australia.

Wagner, D. H., W. C. Mylander, and T. J. Sanders. 1999. *Naval Operations Analysis*. 3rd ed. Annapolis, Maryland: Naval Institute Press.

Wu, H.-F. 2002. Spectral analysis and sonification of simulation data generated in a frequency domain experiment. M.S. Thesis, Department of Operations Research, Naval Postgraduate School, Monterey, California. Available online via <http://theses.nps.navy.mil/02sep_Wu.pdf> [accessed May 21, 2003].

AUTHOR BIOGRAPHIES

SUSAN M. SANCHEZ is a Professor of Operations Research at the Naval Postgraduate School, where she also holds a joint appointment in the Graduate School of Business and Public Policy. She received her B.S. in Industrial and Operations Engineering from the University of Michigan, and her M.S. and Ph.D. in Operations Research from Cornell University. She is a member of INFORMS, DSI, ASA, and ASQ, and is currently the President of the INFORMS College on Simulation and the INFORMS Forum on Women in OR/MS. Her editorial positions are the Simulation Area Editor for the *INFORMS Journal on Computing* and an Associate Editor for *Naval Research Logistics*. Her e-mail and web addresses are <ssanchez@nps.navy.mil> and <http://diana.or.nps.navy.mil/~susan>.

HSIN-FU WU is a Lieutenant in the United States Navy. He graduated from the Naval Postgraduate School in September of 2002 with a M.S. in Operations Research. He received his B.S. in Aerospace Engineering from the University of Kansas. He is a member of INFORMS. He is also a private pilot and is a member of AOPA. LT Wu is a Submarine Warfare Officer in the Navy and is the Combat Systems Officer on the PCU JIMMY CARTER (SSN 23) in Groton, CT. His e-mail is <wuh@carter.navy.mil>.