AD-A083 516 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOO--ETC F/G 9/2
TOWARD VALIDATION OF COMPUTER SIMULATION MODELS IN OPERATIONAL --ETC(U)
DEC 79 J M ARNETT
UNCLASSIFIED AFIT/GOR/SM/79D-1 NL

1 of 2
AD-
A083516

ADA083516

AF INSTITUTE OF TECHNOLOGY

LEVEL II

UNITED STATES AIR FORCE

AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GOR/SM/79D-1

Accession For	
NTIS GRA&I	✓
DDC TAB	
Unannounced	
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or special
A	

TOWARD VALIDATION OF COMPUTER SIMULATION MODELS IN OPERATIONAL TEST AND EVALUATION

THESIS

AFIT/GOR/SM/79D-1

James M. Arnett
Captain USAF

Approved for public release; distribution unlimited

AFIT/GOR/SM/79D-1

TOWARD VALIDATION OF COMPUTER SIMULATION MODELS IN OPERATIONAL TEST AND EVALUATION

THESIS

Presented to the Faculty of the School of Engineering

of the Air Force Institute of Technology

Air University

in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

by

James M. Arnett
Captain USAF

Graduate Operations Research

December 1979

Approved for public release; distribution unlimited

Preface

Validating extremely large and complex computer simulations is an extremely difficult task. As a result, very few such simulations in the Air Force have been subjected to formal validation procedures. Yet, important decisions are often based on the results of those simulations without ever questioning their validity. The need for validation is clear; the methodology is not. I sincerely hope that this thesis provides some insight into how to approach that problem.

I owe so much to so many people for their help and guidance. Dr. Marion Williams of HQAFTEC/OA provided the initial proposal of considering validation in the OT&E context. Major (Ret.) Bob Broderson helped me narrow the scope of the applied part of the research down to a manageable size. His assistance in the early stages of trying to independently implement the Georgia Tech Model was invaluable. Steve Stuk and Mike Tuley of Georgia Tech were incredibly patient in listening to and helping me solve other implementation problems over the telephone. The extensive coverage of validation and experimental design, the areas of the thesis that are probably the most valuable, are a direct result of prompting by my thesis advisor, Lt Col Charles W. McNichols. My reader, Lt Col Tom Clark, forced me (in his own subtle ways) to keep the concept of validation in proper perspective. (No, I can never <u>know</u> how valid a simulation is since I can never really <u>know</u> what the real system's response is in every circumstance.

So how much less tenable is my position when I can't even test the real system?) To these people and many more I give my deepest thanks.

This thesis would have never been completed without the support and encouragement of my family and friends. Through it all, my wife, Pris, responded with an inner strength that makes me proud to be her husband. Her willing sacrifice of her time to help me, her positive attitude when I was down, and her patient endurance of the long hours I spent working on this thesis contributed more than anything else to its successful completion.

Finally, I thank God for the friends I gained, the things I learned, and the strength He gave me throughout this experience called graduate research.

Mike Arnett

Contents

Page

List of Figures

List of Tables

Abstract

The operational test and evaluation (OT&E) of the air launched cruise missile provided the context for considering validation of complex computer simulations in OT&E. The clutter and multipath submodel from TAC ZINGER was used in the research.

Published literature on validation was reviewed and categorized as contributing to one or more areas of validation: philosophies, frameworks, general procedures, and specific methods. Emphasis was placed on validation being a problem-dependent process. The goal of that process is an acceptable level of confidence that the actual and simulated data agree closely enough for an inference about the simulation to be a valid inference about the actual system. Perfect agreement can never be reached for all situations a system could encounter. The cost of accepting a given degree of confidence must be balanced against the cost of obtaining a higher degree.

Experimental design was seen as the key to obtaining the maximum amount of information from a limited number of test runs of the actual system. The full factorial design is much more efficient than the classical vary-one-factor-at-a-time approach, but was rejected for most situations as requiring too many runs. Fractional factorial

designs may be used when higher-order interactions between variables may be assumed to be negligible. In most situations, only a few out of many variables in the system are very important. Fractional factorial, random, supersaturated and group-screening designs are all discussed for use as designs to screen out those important factors.

The following approach to validation was suggested. Use a fractional factorial or other screening design to identify the important variables in the simulation. Use a more complete design to specify parameter/level combinations at which to operate the actual system and the simulation for data comparison. Finally, explicitly incorporate decision analysis in judging the validity of the model based on that comparison and the use to which the simulation will be put.

A fractional factorial design was used to screen out the important factors in the multipath and clutter model. Regression analysis of the data supported the hypothesis that only a few variables were very important, but did not support the assumption of the negligibility of higher-order interactions.

TOWARD VALIDATION OF COMPUTER SIMULATION MODELS IN OPERATIONAL TEST AND EVALUATION

I. Introduction

In a 1969 paper, Clifton Lovell stated that up to that time computer simulation had been used primarily in two areas. First, in the simulation of systems that were conceptual in nature, or in situations conceivable in occurrence, but for which actual experience did not exist. Second, in the simulation of systems where the specifics of the system were reasonably well known, but the cost of physical testing was prohibitive. Lovell maintained, however, that there was an area of overlap between these two that had been neglected. That area was in the use of computer simulations in conjunction with operational evaluations and system effectiveness testing (Lovell, 1969:1).

In the ten years that have elapsed since that paper was written, that area of overlap has become increasingly important as a tool of operational test and evaluation (OT&E). The Air Force OT&E, in particular, is concerned with testing new or modified systems for which little or no actual data on the system's operation exists. Test data can be obtained by field testing, but at a cost that severely limits the amount of testing that can be done. By using a computer simulation of the system's operation, additional data for evaluation can be obtained at a much lower cost and over a much broader range of operational conditions than would be feasible with field testing only.

However, before the simulation can be used as a basis for evaluating the actual system, it must be "validated." In Fishman and Kiviat's words, "The agreement between the behavior of a simulation model and the observed behavior of a real system" must be tested.

> This requires empirical data. If a behavioral equivalence can be established between a simulation model and a real system, we may regard the behavior of the model and the system as being consistent [Fishman and Kiviat, 1967b:v].

In this context, the field testing takes on two roles. It provides data directly for evaluation of the system, and it provides data that can be used to validate the simulation. However, the correspondence of the simulation to the actual system cannot be tested under every possible set of conditions. To attempt to do so would obviously defeat the purpose of using the simulation. A problem thus arises: How does one best go about validating the simulation, given a limited amount of available test resources?

Consideration of a specific application of supplementing field testing with simulation will be useful in attempting to solve this problem.

Finally, the main program accounts for the effects of each engagement on the course of the simulation as a whole (BDM, 1977:6).

The submodels used are a set of existing one-on-one engagement models called the "TAC ZINGER" models. The set includes ten models of surface-to-air missile (SAM) systems, and one anti-aircraft (AAA) model. Input to each model includes information on the nature of the target to be engaged (in this case the ALCM), information on the engaging SAM/AAA system, and information on the environment in which the engagement takes place. Information on the target includes, among other paramters, a flight path description, a radar cross-section or infrared-signature table, and vulnerable presented area tables. (Note that the flight path description is an input. The model is only concerned with the target's external characteristics, not its internal operation.) Data on counter-measures employed by the target, i.e., jamming, is also included. The primary inputs for the engaging site are its location and the firing interval for the defensive weapon. Most of the information characterizing the performance of a specific defensive system is built into the particular model representing it. Input information on the engagement environment includes the effects of factors such as clutter and multipath on the site's ability to track its target. This information is generated by submodels within TAC ZINGER. Output from TAC ZINGER includes an intercept condition code denoting

OT&E of the Air Launched Cruise Missile

The Air Force Test and Evaluation Center (AFTEC) is currently involved in the operational test and evaluation of the Air Launched Cruise Missile (ALCM). From an operational standpoint, major emphasis is being placed on testing and evaluating the survivability of the ALCM in an actively hostile environment. Test launches of the missile through a simulated enemy defensive system can be made to obtain data, but the cost of each launch is on the order of $500,000. To increase the amount of data available for evaluation, a digital simulation of the threat system will be used. At a cost per run in the vicinity of $1,000, the simulation can thus be used as an economical source of data, once it is validated.

The model being used to simulate the interactions between the ALCM and a threat system is called "TAC REPELLER." It was developed by the BDM Corporation for Headquarters Air Force, Studies and Analysis (AF/SA) over the past few years. The model simulates the interactions that transpire when several aircraft simultaneously encounter several ground-based defensive systems. Each individual engagement is simulated in detail by a submodel of one aircraft against one defensive system. The overall program simulates the set of events leading up to the individual engagements. The submodels are then invoked to determine the outcome of each of those engagements.

the basic nature of the engagement results, intercept position and time, probabilities of kill and hit, and various other parameters (BDM, 1977:8).

With so many input parameters to consider, validating the models is a very complex undertaking. As a result, the methodology for validating either TAC REPELLER or TAC ZINGER has not yet been developed. The need for this methodology provided the impetus for this paper.

Obviously, TAC ZINGER must be validated before TAC REPELLER can be. Likewise, the submodels within TAC ZINGER must be validated before TAC ZINGER. In developing a methodology for validating any of these models, a logical first step is thus to consider a submodel within TAC ZINGER. That research should provide some insight into the appropriate methodology for validating TAC ZINGER and REPELLER.

Research Model

The submodel chosen is a set of subroutines which calculate the effects of multipath and clutter on the radar tracking error. The model was developed for AF/SA by the Engineering Experiment Station, Georgia Institute of Technology. For brevity, it will be referred to as the "Georgia Tech Model." The model can be easily adapted to run independent of TAC ZINGER, which facilitates its analysis.

The Georgia Tech Model was developed to upgrade clutter and multipath models in TAC ZINGER, and was incorporated in TAC ZINGER in early 1979. The final report on the project (Zehner and Tuley, 1979) documents the model and the effort to validate it. The validation was based on a statistical comparison of simulated data with existing actual data. The existing data contained information, however, from a set of data points that did not allow a full exploration of the ranges of the variables in the model. For the ranges considered, the method used was inefficient in that there was more data than necessary for some cases and not enough for others. For these reasons, the same approach should not be used to validate TAC ZINGER or REPELLER. The methods of comparison of data might be appropriate, but another way of choosing data points needs to be found. The research for this paper thus focuses on three areas: validation, experimental design as applied to validation, and application of these two concepts to the Georgia Tech Model.

Overview of Report

Chapter II discusses validation. The term is defined and several philosophies of validation are discussed. The literature on validation is reviewed. Chapter III considers experimental design as applied to validation. Appropriate experimental designs are discussed and an approach to using experimental design in validation is presented. Chapter IV describes the procedures followed in conducting the research with the Georgia Tech Model. The model is described in some detail, and a few problems in its documentation are discussed. Chapter V presents the results of the data analysis, and Chapter VI provides a brief summary of the research and conclusions drawn from it.

II. Validation

The need for validation of a computer simulation model is quite obvious when decisions about a multi-billion dollar project like the ALCM will be based on the model's results. As computer simulation has advanced over the past two decades, fairly widespread agreement on the definition of validation has developed. However, despite several attempts, there is still no "accepted" general methodology for validation. As Naylor and Finger (1967:B-92) reported in 1967 and Wright (1972:1286) and Kheir (1976:534) reiterated in 1972 and 1976, "Despite the scope of publication and discussion, validation is still thought to be the most elusive of all the unresolved methodological problems in the social sciences." As recently as 1978, in his excellent review of validation literature, Tytula (1978:5) concludes that "there definitely remains a need for some methodology that can be used to validate simulation models." With such studies as background, this research will work toward an efficient validation procedure for the simulations being used in evaluating the ALCM.

Prior to 1967, very little had been written on the subject of validation. One of the earliest uses of the term validity was in Forrester's Industrial Dynamics, published in 1961 (1961:115). Forrester was dealing with dynamic simulation models, and judged "the validity (or significance) of a model . . . by its suitability for a particular purpose." His ultimate test of validity was whether or not better systems resulted from investigations based on model experimentation. As not all simulations have the goal of resulting in better systems, Forrester's ultimate criterion cannot always be used. On the other hand, his concept of suitability for a practical purpose has had considerable influence on later approaches to validation. His statement that "a model is sound and defendable if it accomplishes what is expected of it" (1961:115), also contributes to the philosophy of validation, but offers little practical guidance on how to put the concept into action.

Definition of Validity

In a 1967 RAND Corporation memorandum, Fishman and Kiviat developed their concept of validation as one of three statistical problems that arise in all computer simulation experiments (Fishman and Kiviat, 1967b:v). As cited earlier, their concept of validation was that it tested "the agreement between the behavior of a simulation model and the observed behavior of a real system." The other two areas are verification and problem analysis. Although not everyone has agreed with their definitions for these terms, it has become generally accepted practice to use "verification" and "validation" in the same general sense as they proposed.

Verification insures that the simulation model behaves as the experimenter intends. This deals with the internal data, structure, and logic of the model. It makes no statement about the behavior of the model as compared to the real world behavior of the system. Among the functions it does perform are tests of independence and randomness of any independent random variables generated, and examination of substructure outputs to determine if they behave properly (Fishman and Kiviat, 1976:11,14). As Van Horn notes (Van Horn, 1969:233), good discussions of verification problems can be found in articles by Conway (Conway, 1963), Fishman and Kiviat (Fishman and Kiviat, 1967a and 1967b), and Naylor (Naylor and Finger, 1967; Naylor, et al., 1969). For the most part, such problems are one step removed from the scope of this research. However, in adapting the Georgia Tech Model to run independently of TAC ZINGER, some verification will be necessary.

Problem analysis deals with the analysis and interpretation of the data generated by the experiments (Fishman and Kiviat, 1976:v). In the context of testing the ALCM, this deals with evaluating the ALCM's survivability based on the data generated by both field testing and simulation. In this light, problem analysis is AFTEC's mission and is beyond the scope of this research.

Validation, then, is in the area between verification and problem analysis. It overlaps verification, but must be accomplished before valid problem analysis can begin.

In 1969, Van Horn, building on Fishman and Kiviat's definition, defined validation as "the process of building an acceptable level of confidence that an inference about a simulation is a correct or valid inference for the actual process" (1969:233). By defining validation in this manner, he addresses three key aspects about validation.

First, it is a process. It cannot be postulated nor done all at once. It takes time to build an acceptable level of confidence in a computer simulation model.

Second, it involves an <u>acceptable</u> <u>degree</u> of confidence in the simulation. As Shannon points out, the results of any simulation can be made as accurate as desired, i.e., the variance can be made as small as desired, provided enough samples are taken (Shannon, 1975:209). However, running simulations costs time and money. A trade-off must be made between the benefits derived from further accuracy and the added cost of obtaining that accuracy.

Third, it emphasizes comparison of inferences rather than actual data. This ties in with Forrester's concept of judging the validity of a model by its suitability or usefulness for a particular purpose. For example, if a decision is to be based on a trend in the simulation data, it would be a waste of time and effort to refine the model to the point that it matches the actual system's output to ten significant digits. This is a different problem than simply dealing with variance, as in the second aspect above. This type of model refinement could require major structural changes, extensive data collection to more accurately define underlying probability distributions, or other conceptual changes in the model.

Thus, using Van Horn's definition, validation is problem-dependent. Simulations are built to accomplish a purpose. Validation builds confidence that the simulation actually accomplishes that purpose. Specific procedures for validation, therefore, can only be determined in light of that purpose.

Although each author tends to define validity in his own words, the great majority of the definitions embrace the concepts proposed by Fishman and Kiviat and Van Horn (Naylor and Finger, 1967:B-93; Naylor, et al., 1967:3); Crabill, 1975:230; Driscoll, 1975:1217; Gilmour, 1973:127; Garrat, 1974:916; Nolan, 1972:1257-1260; Schatzoff, 1975: 252; Schlesinger, et al., 1974:927; Wright, 1972:1287; Shannon, 1975:208; Golub, 1976:701). A somewhat lengthy but operational definition that encompasses the general usage of the term follows: Validation is the process of building an acceptable level of confidence that the simulated data agrees with the real data closely enough that an inference about the simulation is a valid inference about the actual system. It is in this sense that the term is used in this research.

Rationalism, Empiricism, Positive Economics

Underlying the different approaches to validation are three different philosophies: rationalism, empiricism, and positive economics. These will be discussed briefly here; for a more complete discussion, see either Naylor and Finger (Naylor and Finger, 1967:B-93 to B-95) or Shannon (Shannon, 1975:212-215). Rationalism holds that a model or theory is simply a system of logical deductions from a set of unquestionably true premises. Premises of this type are known as *synthetic a priori*. Under this philosophy, validation becomes a matter of determining the *synthetic a priori* upon which the model is based. Once these are found, the model can be considered a valid representation of reality. The problem arises in attempting to find the *synthetic a priori*, if in fact such premises of unquestionable truth exist at all (Naylor and Finger, 1967:B-93 to B-94). Empiricism is at the opposite end of the philosophical spectrum. It insists that no assumption or postulate be used that cannot be empirically verified. Holding strictly to empiricism, the model that uses such an assumption is not valid.

Naylor and Finger point out that in economics, most applications have involved a compromise of both these views. Some assumptions can be validly used with empirical verification. Others are not so obvious or accepted and must be tested.

The third philosophy, positive economics, was proposed by Milton Friedman in 1953 (Friedman, 1953). He argued that validation was not concerned with "the validity of the assumptions on which the model rests . . . but rather on the ability of the model to predict the behavior of the dependent variables which are treated in the model" (Naylor and Finger, 1967:B-94). Friedman's approach came under attack for ignoring the model structure and assumptions and focusing only on its output. By this philosophy, the aspect of simulation that Fishman and Kiviat termed verification is nonessential; only the end product matters.

Frameworks for Validation

As a solution to the problems that arise in applying the three above stated philosophies, Naylor and Finger proposed a three-stage process. Although they thought such an approach may be applicable to simulation in general, they held that it was particularly applicable to models of industrial systems. The process was a deliberate attempt to combine the methodologies of rationalism, empiricism and positive economics. The first stage was to formulate a set of hypotheses to describe the system that are, for all practical purposes, synthetic a priori. These hypotheses may require some further verification, but they form a basis for the structure of the model. The second stage is to empirically verify as many of these assumptions or hypotheses as possible, using the "best available statistical tests" to make this verification. The third stage is to test the model's ability to predict the behavior of the system under study. The emphasis of this stage is on the predictive ability of the model. If historical data is used in building the model, the simulation should be expected to generate results consistent with that data. The real test of validity comes in using the simulation to predict results from different input data, then comparing those results with results generated by the actual system's response to the same input (Naylor and Finger, 1967:B-97). Miller (1974:911) and Garratt (1974:916) also dwell on the importance of predictive validity.

In his critique of Naylor and Finger's proposed methodology, McKenney points out that model validity depends on what the model will be used for (McKenney, 1967:B-10). In their critique of the same paper, Schrank and Holt propose that

> . . . the criterion of the usefulness of the model be adopted as the key to its validation, thereby shifting the emphasis from a conception of its abstract truth or falsity to the question of whether the errors in the model render it too weak to serve the intended purpose [Schrank and Holt, 1967:B-104 to B-105].

Thus Forrester's concept of usefulness again enters into consideration.

Tytula sees Schrank and Holt as implicitly formulating the validation process as a decision problem (Tytula, 1978:6). Van Horn explicitly described it in the same way. The problem was "to balance the cost of each action against the value of increased information about the validity of an insight" (Van Horn, 1969:233).

Nolan (1972) also attempted to set down a general framework for verifying/validating simulation models. Although he used his own terminology in describing his procedure, close inspection reveals his problem is basically a reorganization of the concepts already mentioned here and does not add anything new to these discussions (Tytula, 1978:8).

A somewhat different approach to validation has been taken by Hermann (Hermann, 1967:220-224). He developed five validity criteria which, although developed in the context of models of international politics, are applicable in other fields as well.

The first criterion is internal validity. This type of validity is established during the verification of the model. Internal validity has to do with the variance between replications of a simulation. Assume each run begins with identical initial parameter values, and any exogenous variable introduced during a run is the same for each run made. If there is variance among the results of the runs that can be attributed to extraneous factors rather than specified factors in the simulation, then the internal validity is low.

The second criterion is face validity, the surface or initial impression of a simulation's realism. That is, does the output "look right" to someone familiar with the system? Face validity is almost entirely subjective. Therefore, it is of most value during the initial stages of model verification, the time that gross irregularities are most apt to arise. This is not to detract from the importance or value of face validity. Having someone familiar with the system review the output before subjecting it to rigorous statistical testing can save a great deal of time and energy.

The third criterion, variable-parameter validity, involves comparisons of the simulation's parameters with their corresponding values in the real system. Sensitivity analysis is one feature of variable-parameter validity, and some form of variable-parameter validity is the most common approach to statistical validation. The approach offers the advantage of being able to isolate individual parameter's contributions to the model's correspondent to the real world. A common misconception in applying this technique is that only one parameter should be varied at a time. As will be discussed in more detail in the next chapter, such a procedure is unnecessary and inefficient. Furthermore it cannot account for the synergistic effect that two or more variables, when acting together, may have on the results. A drawback to variable-parameters validity is that it does not directly consider relationships between parameters. The contribution of these relationships to the overall validity of the model is discussed in the fifth approach.

Event validity, the fourth approach, basically addresses the issue of isomorphism. That is, how closely does the model have to resemble the actual system? To what degree do elements in the model have to correspond to elements in the real system? Furthermore, at what level of generality should events in the model and in the real world be compared? Is it enough to simply know that the defensive system destroyed the ALCM, or does every instant of interaction need to be simulated and reported, or is there an appropriate approach somewhere between the two. Answering these questions satisfactorily will build confidence that the simulation is accomplishing its intended purpose.

In the fifth approach, hypothesis validity, the criterion is the correspondence of hypothesized relationships between their counterparts in the model. Two kinds of relationships can be used in hypothesis validity: first, the programmed relationships which are an integral part of the model; and second, relationships that exist, but are independent of the programmed relationships. The first type may be stated as researchable hypotheses or they may be empirically derived. The second type are typically more important in the social sciences; data-oriented simulations for other applications tend to build into the model as many relationships as can be found. A final note on this fifth approach: a set of relationships that involve a variety of different model components should be considered. This will provide a broader base upon which to judge the validity of the model.

Hermann goes on in his paper to state that attempting to satisfy these criteria builds confidence in the overall validity of the model. He further points out that validity is always a matter of degree (Hermann, 1967:225). Thus, his approach considers the important aspects of validation, and his five criteria provide insight into the many connotations of the word validity.

The authors of the preceding approaches proposed a number of different methods of comparing simulation results with real world results. Naylor and Finger point out that "management scientists and economists have, more often than not, restricted themselves to purely graphical (as opposed to statistical) techniques of 'goodness of fit'" (Naylor and Finger, 1967:B-97). They go on to list what they considered to be the more important statistical tests available for testing goodness of fit. Those tests were: analysis of variance, chi-square test, factor analysis, Kolmogorov-Smirnov test, nonparametric tests, regression analysis, spectral analysis, and Theil's Inequality Coefficient (Naylor and Finger, 1967:B-98 to B-99). They basically describe what each test does and when it is useful. They deferred further discussion to appropriate procedures which describe the given test in detail. Van Horn refers to these eight, and adds spectral analysis and the "Turing" test--have people directly involved in the actual process compare simulated input and output data. If they cannot discriminate between the two, the simulation is assumed to be valid (Van Horn, 1969:240-242). A few other techniques have been proposed in more recent literature.

The approaches just described provide frameworks for validation, from which procedures can be developed for specific applications. The authors did not claim to be presenting specific methods that could be used to validate <u>any</u> simulation--they were discussing validation in general. A few authors have, however, published "general validation procedures" they believed could be applied to amost any validation effort.

General Validation Procedures

Gilmour, in developing his generalized validation procedure (Gilmour, 1973:127-131), divides validity into two parts: design validity and output validity. Design validity is a combination of what has been previously referred to as verification and face validity. Output validity "examines and establishes the requisite quality of the model's endogenous data streams" (Gilmour, 1973:127). In his view, the establishment of design validity is problem dependent, but a general procedure can be used to establish output validity in any model. His procedure thus does not deal with establishing design validity, although he acknowledges that it is "important as cost-benefit considerations limit the more comprehensive output validity testing to the several key variables" (Gilmour, 1973:127). He implies that demonstrating design validity will establish what those key variables are. He does not address the issue of how they are established, a significant problem in complicated models with many variables. These problems aside, his procedure for establishing output validity is straightforward and would indeed develop confidence in the model.

His procedure has three parts. It employs a set of statistical tests to establish the stability of the model over time, measure the predictive ability of the model, and perform sensitivity analysis on the major assumptions of the model. Fourteen statistical tests are listed, with the recommendation to use as many as possible, subject to cost constraints and the applicability of individual tests. He proposes, finally, calculating an index of validity for each of the three classes of output validity (stability, predictive ability, and sensitivity to assumptions). This index is based on weighting the value of each statistical test by an appropriate value. The weight he proposes is the inverse of the number of underlying assumptions upon which a test is based that are violated when that test is applied to the model. The index is calculated from the following equation:

$$V_j = \frac{\sum_{i=1}^{n} \left(\frac{P_i}{n_i+1}\right)}{\sum_{i=1}^{n} \left(\frac{1}{A_i+1}\right)} \tag{1}$$

where V_j is the index of validity of the j^{th} class, n is the number of tests applied to that class, P_i is the percentage of favorable results using the i^{th} test, and A_i is the number of assumptions violated in using the i^{th} test.

Gilmour suggests this index as a "heuristic tool for the simulation experimenter," and cautions against using it to infer "absolute" validity (Gilmour, 1973:129). He does believe it can be useful in comparing his three classes of output validity, in evaluating the effect on validity of changes in variables, and in comparing the validity of two or more models (Gilmour, 1973:130).

Gilmour's approach seems basically sound. The assumption of having an existing model that has passed design validity tests is, for many applications, appropriate. The present research using the Georgia Tech Model is an example. The use of several tests of hypothesis helps build confidence in the validity of the model. Performing sensitivity analysis on the major assumptions of the model can establish a level of confidence in the model's ability to perform correctly when those assumptions cannot be met exactly. On the negative side, he provides no suggestions as to how to identify the "key" variables to be examined for output validity. Also, until proven in other contexts, the ability of his "validity index" to truly measure validity should remain suspect. Although the article was published in 1973, the only reference to it found in the literature was in the 1974 volume of <u>Computing Reviews</u>. Perhaps this is because the article was published in an Australian journal. In any event, it should be given further consideration in future applications.

A standard procedure for validation based on sensitivity analysis is proposed by Miller (1974). He describes his approach as depending "only on the magnitude of uncertainties in the data, the sensitivity of the results to errors, and comparison of the accumulated uncertainty to a user-defined limit." He maintains that the results of such testing will point to specific areas that need refinement (Miller, 1974:911). Tytula notes that Miller's procedure makes it difficult to assess the effects that model structure changes have on the validity of the model. This problem may also explain why other validation approaches based on sensitivity analysis have not been widely used (Tytula, 1978:14).

Schlesinger, et al., (1974:928-933) propose a rather concise "standard" set of procedures for verification and validation of a model. Their proposal falls between the frameworks discussed earlier and the "general" procedures above. They do not add anything new to the philosophies or concepts presented in the frameworks, but they stop short of advocating a specific methodology as proposed by Miller and Gilmour.

They first require the model be verified, emphasizing the use <u>and documentation</u> of (1) numerical test cases that check all major facets of the model and

(2) logic flow tests that verify the correct implementation of all control statements in the program.

Their next step describes four aspects of establishing the "reasonableness" (face validity, as described earlier) of the model: continuity, consistency, degeneracy, and internal validity tests. Continuity insures that small changes in input parameters result in small changes in output, unless there is specific justification for expecting abrupt changes. Consistency requires the model produce essentially similar results when essentially similar cases are run, even though they are input with differing combinations of descriptive parameters. Degeneracy insures that when parameters are chosen which eliminate the effect of a feature in the model, the model does in fact react as if that feature were not there. Internal validity tests are implemented within the program to check for logically absurd conditions, such as an aircraft flying below the surface of the ground.

The validation procedure requires using some quantitative measure of deviation of simulated from actual data, but leaves the determination of that measure up to the analyst. The matching of relatively more important parameters should be given more emphasis when comparing the real and simulated data. The validation procedure must further include a qualitative discussion of significant deviations (Schlesinger:927). This part should include a discussion of not just the raw data, but also inferences from the data appropriate to the specific situation, as suggested by Van Horn.

The fourth step of the procedure is to insure the model is not used outside its "domain of applicability." That is, the model must not be used in situations that violate the assumptions on which it is based.

Schlesinger, et al., make one final point not discovered elsewhere in the literature: the importance of exercising control over changes made in a validated simulation model. They refer to a model that has been successfully verified and validated as "certified." Once a model has been certified, any changes to that model must be accompanied by appropriate testing to insure that validity is maintained. A copy of the exact computer code used, as well as documentation of the verification and validation procedures followed must be maintained for future reference. Experimentation may only be done with a certified model (Schlesinger, et al., 1974:928).

Specific Methods

Tytula (1978:9) divides the specific methods that have been used in validating simulations into five general categories: judgmental comparisons, hypothesis testing, spectral analysis, sensitivity analysis, and indices of performance. Most examples found used more than one approach, most often the judgmental combined with one or more of the others.

Judgmental Comparisons. Judgmental comparisons involve graphical analysis and comparison of common properties of the real system and the model. Tytula (1978:9) defends its practicality, since the human eye can recognize all sorts of patterns and relationships that are difficult or impossible to discern by quantitative techniques. In the literature, the approach has been used to build confidence in the model by closely checking the model's logic and transformation into a computer program, then visually comparing the simulated and actual data. Examples of its use include the validation effort of the Georgia Tech Model now under consideration, (Zehner and Tuley, 1979); validation of a police patrol model (Crabill, 1975); missile system simulations (Driscoll and Stockdale, 1975; Golub, 1976; Kheir, 1976); and a ballistic missile defense system (Kosovac and Shortle, 1976). Tytula points out two significant drawbacks to the approach: the scale used in making the comparison can have a significant effect on the judgment of how good the comparison is, and there is no way to assess the impact of errors in judgment (Tytula, 1978:10).

Hypothesis Testing. The second category, hypotheses testing, is widely used also, and relates back to the developers of the frameworks discussed earlier. These techniques are applicable when comparing output streams of data points that are independent with respect to time. These techniques can vary greatly in their assumptions, purposes, and methodology. Included are parametric tests such as the F and Student's t tests; nonparametric tests like the Mann-Whitney test of means; tests to determine an underlying probability distribution, such as the Kolmogorov-Smirnov test; and tests to determine the degree of relationships between two or more variables, such as correlation analysis (Shannon, 1975:228). Garratt discusses multivariate analysis of variance (MANOVA), permutation, and nonparametric ranking methods (Garratt, 1974:917-921). All of these tests are based on specifying an acceptable level of confidence (e.g., 95%) that the test will show the model valid if the simulated data agrees with the actual data. Such a test makes no statement about the probability of accepting the model when the two sets of data do not agree. In most instances, this information would be more valuable. Rejecting the validity of an actually acceptable model would not usually result in as serious consequences as accepting the validity of one that is not valid, and then basing a decision on its erroneous output. The difficulty in establishing this second probability, known as the probability of making a Type II error, is well established for most tests (Mendenhall and Scheaffer, 1973:338). It involves constructing a family of alternative data streams against which the simulated data stream can be compared. For most simulation problems, this task is especially difficult. From this development, Tytula concludes that "hypothesis testing methods are usually not applicable because they tend to examine the wrong issue" (Tytula, 1978:12).

Spectral Analysis. To solve the problem of comparing streams of autocorrelated data, Fishman and Kiviat applied the technique of spectral analysis to computer simulation (Fishman and Kiviat, 1967a:526-527). Spectral analysis does overcome the autocorrelation problem, but in the process develops several other serious advantages (Van Horn, 1969; Watts, 1969; Howrey and Kelejian, 1969; Tytula, 1968). The simulations under consideration here produce data that can be tested using non-time-series procedures, so the technique will not be further discussed here.

Sensitivity Analysis. The fourth category is sensitivity analysis. Van Horn proposed using it in place of strict empirical testing of underlying hypotheses, assumptions, and parameters. Such empirical testing requires a lot of data, and so is very expensive. Van Horn argues that an insight gained from a simulation is normally valid over a range of parameter values and usually does not depend on a specific distribution. Sensitivity testing can then be used to "establish the set of distribution and parameter values for which a set of insights is relevant" (Van Horn, 1969:234). The requirement for and cost of empirically testing assumptions could thus be reduced. Miller's use of sensitivity analysis has already been discussed, as has Tytula's concern about the approach's difficulty in analyzing the effects of model structure changes.

Performance Indices. Tytula's final category of validation methods are those that use some index of performance or fit. Theil's inequality coefficient (TIC) has been the primary index used, although Gilmour's index of validity may fit in this category. Tytula (1968:14) cites the form of TIC most commonly used (e.g., Naylor and Finger, 1967:B-99; Kheir, 1976:536) as

$$\mathrm{TIC} = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^{n} (P_i - A_i)^2}}{\sqrt{\frac{1}{n} \sum_{i=1}^{n} P_i^2} + \sqrt{\frac{1}{n} \sum_{i=1}^{n} A_i^2}} \tag{2}$$

where n is the number of elements in a sequence, and P_i and A_i are the prediction and realization, respectively, of the $i^{\underline{th}}$ element. The value of TIC runs between 0 and 1, with 0 being perfect agreement between the prediction and the realization, and 1 corresponding to no agreement whatsoever (Theil, 1961 or 1970).

Naylor and Finger (1967:B-99) cited the use of TIC by a number of economists to validate simulations with econometric models. In that context, TIC provided an index which measured the degree to which a simulation's predictions agree with actual data. Kheir (1976:536) used TIC in evaluating missile-systems simulations.

Schrank and Holt (1967:B-105 to B-106) propose using a somewhat different form of Theil's Inequality Coefficient (Theil, 1966:28):

$$U = \frac{\sum_{i=1}^{T} (P_i - A_i)^2}{\sum_{i=1}^{T} A_i^2} \tag{3}$$

where T is the number of time periods, corresponding to n above; and P_i and A_i are defined in the same manner as above. Using this form, they propose an overall coefficient of performance, C, for N forecasted variables:

$$C = \sum_{j=1}^{N} w_i U_j^2 \tag{4}$$

where w_j is a weight indicating the importance of the intended application that is attached to forecasting the $j^{\underline{th}}$ variable. Tytula (1978:6) sees Schrank and Holt as formulating validation as a decision problem.

Tytula notes that with the exception of Schrank and Holt's formulation, there is no apparent connection between TIC and the implications of actions taken on the basis of the model output. As a result, he concludes "it is not possible to pick a rational value for TIC which can be used as a criterion for deciding that a model is valid"

(Tytula, 1978:14-15). He does, however, further note TIC's usefulness in ranking alternative models of a process.

As a result of his analysis of previous validation work, Tytula (1978:15) concludes that none of those methods were without significant pitfalls:

> The most important of these shortcomings are the inability to handle the autocorrelation of the simulation output variables, concentration on the wrong issue, and difficulty in transforming the measure of disagreement between simulated and actual results into some meaningful set of consequences.

He then develops an approach that he proposes will overcome those problems. Although the issue of autocorrelation is not relevant to the present study, the other two issues are. The approach is based on modern decision theory, as opposed to classical hypothesis testing and is briefly summarized below.

Decision Analysis. As in hypothesis testing, one of two decisions will be made based on a comparison of simulated with actual data. Either the simulation will be accepted as valid, or it will be rejected. Using decision theory, the problem can be formulated as shown in Figure 1, where a square represents a decision and a circle represents a chance event.

Even when a simulation is accepted as being a valid representation of the real system, there still will virtually always be some difference between the simulated and actual data. Since both data streams can be modeled as stochastic processes, the error between them is also a

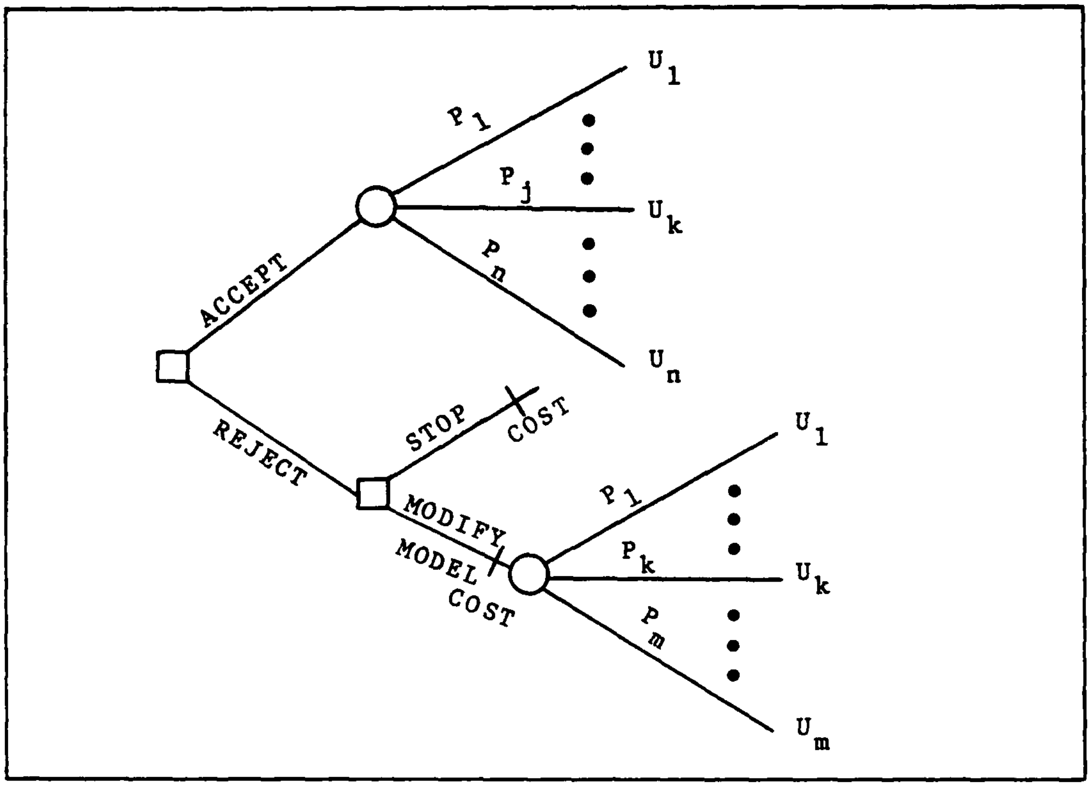

Figure 1. Generalized Validation Decision Model (Tytula, 1978:78)

stochastic process, and has a probability distribution. The probabilities of each value of that error distribution are the P_i shown in Figure 1, i=1 to N, the number of possible values the error distribution can take on. Associated with each of these N values is some utility, u_i, based on the specific purpose for which the simulation is to be used. Knowing the probability and utility associated with each possible value of the error distribution, the overall expected utility of the decision to accept the simulation as valid, $E(U_v)$, can be calculated from Equation 5.

$$E(U_v) = \sum_{i=1}^{N} P_i u_i \tag{5}$$

If the simulation's validity is rejected, another decision must be made: either modify the model or stop the evaluation. Again either decision involves a cost. If the model is modified, a new error distribution will have to be determined. Once utilities for each value of the new error distribution are established, the expected utility of initially rejecting the validity of the model can be calculated. If the expected utility of accepting the simulation's validity is higher than the expected utility of rejecting it, the model can be considered valid, but <u>only</u> for the <u>specific purpose</u> for which the individual utilities were calculated (Tytula, 1978:77-78).

The first step in applying this procedure is determining the error distribution. Tytula develops the approach in the context of a missile system simulation, and shows that he is dealing with time-series data (Tytula, 1978:18). As a result, both the actual and simulated data streams can be represented by stochastic processes. The data from a test flight is viewed as one sample realization of the underlying stochastic process. Tytula uses a linear Box-Jenkins model to transform the test data into a model which has the same stochastic properties as the underlying process. This model is then used to generate additional sample realizations, from which the marginal distributions of the flight test data are calculated. Using Monte Carlo techniques, the simulation is replicated to generate the marginal distributions of the simulated data. With the marginal distributions of both sets of data known, the conditional distribution of the error can be calculated (Tytula, 1978: 27,30-31,77).

After the error distribution is found, the utility of each possible value of that error must be determined. As mentioned, the measure of utility used will depend on what the simulation is to be used for. For many applications, cost is a convenient and appropriate measure of utility. If cost is used, the option with the lowest expected cost will have the highest expected utility. Once the measure of utility is determined, the problem of the number of values that measure can take on must be dealt with. Fortunately, most purposes will allow setting acceptable ranges of error for which one value of utility will hold. When this can be done, it will greatly reduce the number of utility values to be determined. Even so, the task of assigning a utility value to each error range can be formidable (Tytula, 1978:78-79).

Thus, Tytula's approach provides appropriate methodology for dealing with the three pitfalls of previous methods. However, for simulations that do not result in time-series data, the approach's applicability is not as direct. Some method of transforming the error into a probability distribution is still needed. A Box-Jenkins model works well for time-series data, but cannot be applied to time-dependent data. Had Tytula's approach been discovered earlier in this research, an attempt to apply it to the Georgia Tech Model would have been made.

Although Tytula (1978:88) finds fault with classical hypothesis testing and recommends his own procedure, he does not completely rule out using hypothesis testing. Until his methodology is extrapolated to non-time-series data, classical testing will have to be used.

Experimental Design in Validation

In addition to the three shortcomings Tytula identified in published validation efforts, very few of those efforts discuss one other important consideration: what to do about not being able to compare the simulation and the real system at every possible combination of parameter values. Most, if the subject is even mentioned, simply choose reasonable values to run test cases with. They do not explicitly consider what combination of cases will provide the most usable information for a given set of runs. The use of experimental design to accomplish this goal in computer simulation experiments with an assumedly valid model is well-developed and documented. Naylor (1969), Naylor, et al. (1968), Kleijnen (1975), Shannon (1975), and others have published books on or including the topic. However, only one reference was found in the literature that specifically applied experimental design techniques to develop an efficient validation design (Schatzoff and Tillman, 1975).

The problem of too many combinations of parameter values has two parts: the number of parameters and the number of values each parameter may have. Both of these problems will be discussed in detail in the next chapter, but the former deserves some introductory remarks here.

In most systems, only a relatively few parameters have a significant effect on the output of the system. If this holds for a system under consideration, then experimental design techniques could be used to identify those important parameters. Validation of the simulation can then concentrate on testing the agreement of the effects of those parameters between the simulation and the real system. This concept of focusing on key variables is suggested by Gilmour (1973:127) and Kheir (1976:537). However, they offer no suggestions for identifying those variables. Typically, people familiar with the system are consulted to determine what they consider the important variables to be. In many cases this method by itself is acceptable. In others, though, such experts are not available or else the effects of the individual parameters are not well enough known to allow its use. Such cases require a more rigorous approach. Such an approach is discussed in the next chapter. In application, a combination of the two approaches is probably most efficient.

In light of this lengthy discussion of validation philosophies, frameworks and applied methodologies, an approach was developed to apply to the Georgia Tech Model. No single approach found could be directly applied, but several contributed key ideas that could be combined into a specific methodology.

First, the model, as implemented independent of TAC ZINGER, had to be verified as operating properly. The model had previously undergone validation testing and had been used in TAC ZINGER for several months with no apparent problems. Therefore, a full-scale verification process was not expected to be necessary. However, test data runs were made to insure at least the face validity of the model's output. Such procedures are necessary whenever an existing model is implemented under new circumstances, but they are not always documented. A later chapter describes the verification accomplished on the Georgia Tech Model as implemented for this research.

Second, experimental design techniques were used to determine the "key" input variables. Consider, as Schatzoff and Tillman (1975:252) suggest, the simulation as a "black box." That is, only the values of the input and output variables matter, not the transformation process between them. By systematically varying the input variables and analyzing their effects on the output variables, the input variables with the greatest effect on the output variables were identified. Two factors had to be considered to accomplish this: the magnitude of the change in the output variable and the magnitude of the change in the input variable. To account for scale differences, the magnitudes of the changes in the input variables were "standardized" as their maximum expected ranges. In this way, the changes in the output variable due to the change in each input variable could be directly compared.

The research with the Georgia Tech Model stopped at this point. A reasonable approach to completing the validation process is developed in the next chapter.

III. Experimental Design in Validation

To maximize the amount of information obtainable from any experiment, a systematic means of specifying input parameter values needs to be used. The principles of accomplishing this are part of the field of experimental design. Experimental design techniques are applicable to virtually any field of experimental research, and have expanded greatly since the publication of Fisher's The Design of Experiments in 1935. Because the field is so broad, much of the material it encompasses is not relevant to a particular application. Such is the case with computer simulation, and in particular, validation of simulations. Naylor, Burdick and Sasser point out the difficulty in separating the material relevant to simulation from the overall body of knowledge of experimental design (Naylor, et al., 1969:3). This chapter attempts to deal with only that information relevant to simulation validation.

To minimize the confusion in discussing this complicated subject, some "standard" usage of terms needs to be made at the onset. Because experimental design techniques have been developed somewhat independently in so many fields, terms appropriate to each field's particular applications have been used. As a result, the terms input variables or parameters, factors, treatments, independent variables, and exogenous variables are all commonly used to refer to the same thing. Similarly, output variables and endogenous variables are used interchangeably. Here, factor and response or response variable will most often be used. The values that each factor may hold are called levels. The entire range of possible responses defines the response surface.

As noted in the previous chapter, the topic of optimum factor-level combinations in simulation validation is virtually ignored by published material. Yet, considerable savings and increased confidence in the simulation's validity over a useful range of conditions should result from using appropriate experimental design techniques in that application. What techniques are appropriate depends on the objective of the experiment. In a broad sense, experiments are run for one of two reasons. They may be run to either generally explore the response surface, or they may be run to determine the optimal point on the response surface over some region of interest in the factor space (Hunter and Naylor, 1969:41). The present research falls in the first category, thus only experimental designs applicable to exploratory experiments will be considered.

Once the experimental data are collected, some means of data analysis must be employed. In simulation, two methods are commonly used: analysis of variance (ANOVA) techniques and regression analysis. ANOVA has been more frequently used when qualitative factors are present, and regression analysis when all of the factors are quantitative. However, either can be used to analyze quantitative and/or qualitative data. Due to limitations of available computerized ANOVA routines, regression analysis will be used in this research.

Designs for Exploratory Experiments

As the field of experimental design has developed, various schemes have been used to methodically explore all or portions of a response surface. However, many of these techniques are not of interest in experiments with computer simulations. Techniques for randomization and blocking, for instance, were developed because of the incomplete control of experimental conditions in typical industrial and agricultural experiments (Kleijnen, 1975:287). Computer simulation experiments, however, offer the degree of control necessary to use other, more efficient designs. Hunter and Naylor discuss four types of designs applicable to exploratory experiments with simulations: full factorial, fractional factorial, rotatable and response surface designs (Hunter and Naylor, 1969:43-53).

Each technique requires a different number of design points to describe the response surface, where each design point is one particular combination of the levels of the input parameters. The response surface could be estimated from only one run of the simulation at each design point. However, most simulations have random processes in them.

The results from only one sample (i.e., one run) cannot be considered a good estimate of the actual response at that design point. To increase the accuracy of the estimated response, several runs (replications) are made for each design point. This leads to the primary problem in applying any of the above techniques: the expense incurred in obtaining the data needed to describe the response surface. The applicability to validation of the four above techniques, plus the "classical" one-factor-at-a-time approach are discussed below.

<u>The One-Factor-at-a-Time Approach</u>. Prior to the publication of Fisher's <u>The Design of Experiments</u>, general practice in experiments was to vary only one factor each run of an experiment (Schatzoff and Tillman, 1975:254). This technique is very inefficient; it requires a large number of runs for the information it provides. The experiment is replicated at each level of interest for each factor to estimate the mean response of the system. As one factor is varied, all others are held constant. Table I shows the approach for two factors at two levels each, each with 6 replications. In all, 24 observations are required.

Once the mean response of the system for each level of each factor has been found, the overall mean response of the system can be determined. The difference between the overall mean response and the mean response for a factor at a particular level is an estimate of the individual or <u>main effect</u> of the factor at that level. This estimate may not

TABLE I

THE ONE-FACTOR-AT-A-TIME APPROACH

Factor A		Factor B	
Level 1	Level 2	Level 1	Level 2
X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X

be accurate, however, due to interactions between variables. Simply put, interactions between variables exist when the system's response to one factor is not the same for all levels of the other variable(s). When interactions exist, the main effect of a variable must consider the system's response at more than one level of the other variables. Interactions between at least two factors are present in most systems, so in most cases, the one-factor-at-a-time method is deficient in that it provides neither estimates of the interactions nor accurate estimates of the main effects.

Full Factorial Designs. To remedy the inadequacies of the one-factor-at-a-time approach, full factorial designs combine all levels of a factor with all levels of all other factors. The mean response of a factor at a given level is calculated as the mean response over all the levels of every other factor, all possible interactions may be calculated. With the one-factor-at-a-time approach, each design point required a number of replications, r, to estimate the mean response for a given level of a factor. In a full factorial design, more than one design point contains a given factor at a specific level. By accepting a slight loss in the number of degrees of freedom for the error term the replications for that factor and level can be divided up among these design points that contain the factor and level (Table II). This reduces the accuracy of the estimated mean response at any design point. However, this is offset by the fact that the mean response for a given factor and level is now properly calculated over all the levels of the other factors. Thus, the full factorial design can achieve about the same accuracy as the one-factor-at-a-time approach with fewer replications per design point.

TABLE II

FULL FACTORIAL APPROACH

		Factor B	
		Level 1	Level 2
Factor A	Level 1	X X X	X X X
	Level 2	X X X	X X X

This efficiency and completeness in estimating the effects of all the variables has its price, however. For k factors each with n levels full factorial designs require n^k design points to completely describe the response surface, and is called an n^k design. If r replications are made at each design point, rn^k total runs of the experiment will be necessary. To illustrate the expense of running a full factorial design, consider a very small simulation that has only seven factors, each of which will be set at two levels. If each run takes only 15 seconds of computer time and each design point is replicated 10 times, a full factorial design would require 5-1/3 hours of computer time (Hunter and Naylor, 1969:44). Thus, even with simulations which have a very few variables, full factorials can require a large amount of computer time. For more than a few variables, the time required would prohibit the use of this design.

Fractional Factorial Design. Fractional factorial designs are incomplete, or fractions of, full factorial designs. Observations are taken only at certain design points instead of at all possible combinations of levels. By choosing these design points properly, main effects and some interactions can be estimated with only a fraction of the runs required for a full factorial design. These estimates are not as accurate as can be obtained from a full factorial design, but they can be accurate enough to determine which main effects and interactions are the most important. Fractional factorial designs have found their greatest use in this area (Hunter and Naylor, 1969:45). A fractional factorial design was used for this purpose in this research, and will be discussed in the next chapter.

Any time a full factorial design is not used, estimates of effects will be in error because of confounding of effects. That is, the design cannot distinguish between two or more different effects. As an example, consider a fractional factorial design that confounds the main effect of factor A with the interaction between factors B and C (a two-factor or <u>second-order</u> interaction). When the effects are estimated, both will have the same value. What is actually measured is the sum of the main effect of factor A and the interaction between factors B and C. The design does not provide enough information to separate them into their individual effects. Effects like these that are confounded with one another are called each other's <u>aliases</u>.

The usefulness of fractional factorial designs is based on the premise that higher-order (usually three-factor or greater) interactions are zero or small enough to consider negligible. This assumption is often valid under the "conditions of smoothness and similarity commonly encountered (Box and Hunter, 1961a:311). Under this assumption, main effects and two-factor interactions can be estimated fairly accurately by choosing a design that confounds them with third- or higher-order interactions only.

A common approach in using fractional factorials to screen for important factors is to use only two levels of each factor. For quantitative applications the two levels correspond to a high and a low value. For qualitative applications they may represent, for example, an on or an off, or a yes or a no answer. Using only two levels allows a great reduction in the number of runs without, under appropriate conditions, seriously degrading the validity of the results of the screening process.

To illustrate the reduction in the number of runs, consider a full factorial design with 8 factors with 4 levels each or a 4^8 design. The 1/16 fraction of that design requires $(1/16)(4^8) = 4069$ design points, and is called a 4^{8-2} fractional factorial design. If only two levels can be used to map the response surface only $(1/15)(2^8)=16$ design points are needed. This design is denoted as a 2^{8-4} fractional factorial design. But, again, nothing is free, and to achieve this dramatic decrease in the number of design points, the ability to estimate non-linear responses to the factors has to be given up.

As long as the response can be assumed to increase or decrease monotonically between the high and low levels, the two level approach to screening is valid. If the response is linear, as shown in Figure 2a, no information is lost by using only two levels of a factor. In Figure 2b, the response is not linear. But can be reasonably represented by a line through the response at the high and low levels of the factor. Figure 2c shows a case for which the two level approach is invalid. In this case, the response is the same at the high and low levels of the factor, and is much greater between those values. In this situation, the main effect of the factor would be estimated as zero, indicating the factor has no influence on the response. In

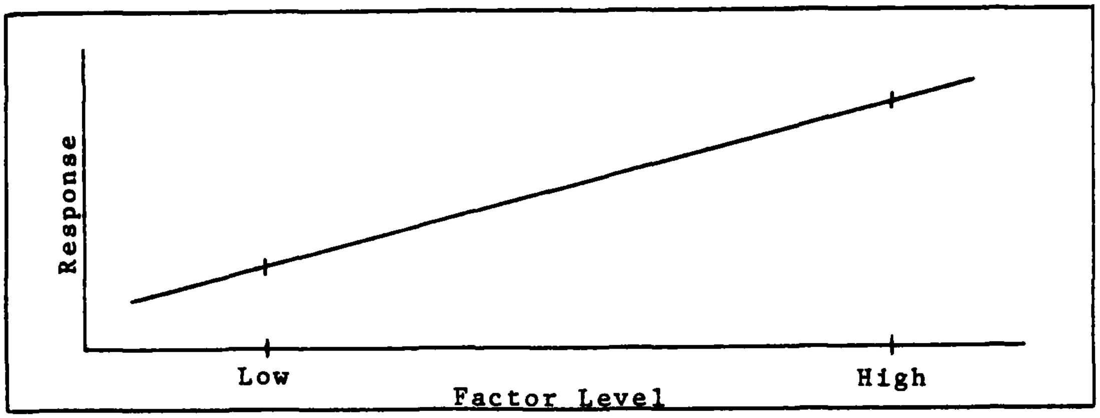

Figure 2a. Linear Response to a Factor

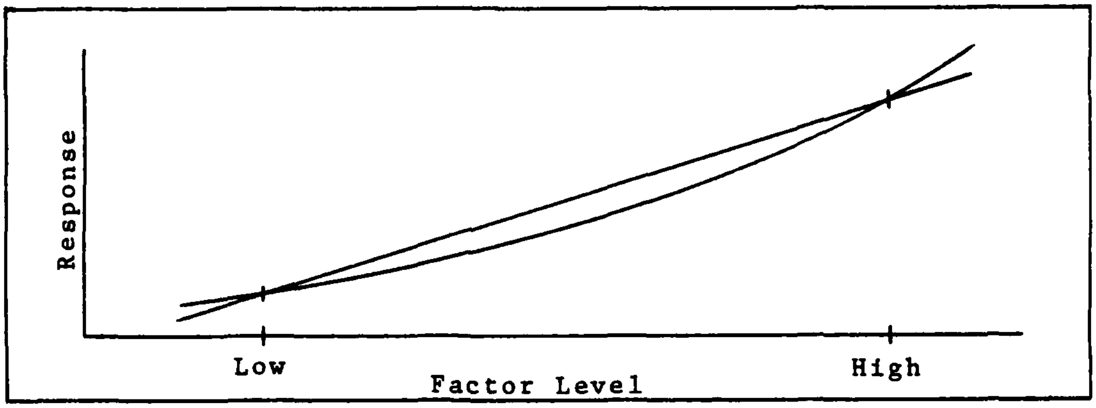

Figure 2b. Reasonable Linear Representation of a Nonlinear Response

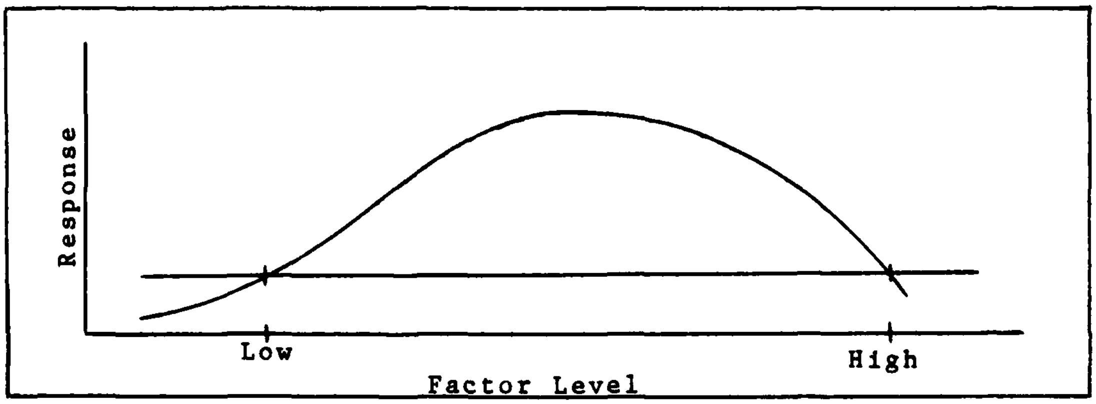

Figure 2c. Unreasonable Linear Approximation to a Response

actuality, the factor has a large influence on the response, but at levels not considered.

To account for severe nonlinear behavior such as this, one or more additional observations must be taken between the two points already observed. In actual situations, a perfectly linear response like that in Figure 2a would be extremely rare. For many, perhaps most, applications, the second case would be the expected situation. This case held for all of the factors considered in this research. The third case is possible in many instances, and care must be taken to insure it is properly dealt with.

Before proceeding further, a visual example of a fractional factorial design will be useful. Table III shows the construction of a 2^{4-1} fractional factorial design. The design has four factors, each with two levels, and is a one-half fraction of a full factorial. It requires $2^{4-1} = 2^2 = 8$ runs, as shown. A "+" indicates the factor is set at its high level; a"-" indicates it's at its low level. The levels of factors 1, 2, and 3 are shown in <u>standard order</u>: column n consists of alternative groups of 2^{n-1} minus signs followed by 2^{n-1} plus signs. The level of factor 4 in any row is the product of the signs of factors 1, 2, and 3 in that row. The relationship $\vec{1}\vec{2}\vec{3}=\vec{4}$ is shorthand for saying that the main effect of 4 will be measured as, or confounded with, the three-factor interaction between factors 1, 2, and 3.

TABLE III

A 2^{4-1} FRACTIONAL FACTORIAL DESIGN
(Box and Hunter, 1961a:314)

	Design Matrix				Observation
Run	1	2	3	123 = 4	Y
1	-	-	-	-	8.7
2	+	-	-	+	15.1
3	-	+	-	+	9.7
4	+	+	-	-	11.3
5	-	-	+	+	14.7
6	+	-	+	-	22.3
7	-	+	+	-	16.1
8	+	+	+	+	22.1

Other effects are also confounded, and can be determined using the defining relation, which is "the key to all the relationships between the effects" (Box and Hunter, 1961a:315). The defining relation is determined from the set of interactions that result in a column of all plus signs, known as $\vec{I}$. Notice that any column multiplied by itself results in $\vec{I}$, and that $\vec{I}$ times any column is that column. (Note that this is not vector multiplication. In this usage, each element in a vector is multiplied by its corresponding element in the second vector, the result is another vector, not a scalar.) For the design in Table III, multiplying the relationship $\vec{4}=\vec{1}\vec{2}\vec{3}$ by $\vec{4}$ results in $\vec{I}=\vec{1}\vec{2}\vec{3}\vec{4}$. Any product of effects that results in $\vec{I}$ is called a generator of the design. In this example $\vec{1}\vec{2}\vec{3}\vec{4}=\vec{I}$ is the only generator and is also the defining relation. In the general case, the defining relation is composed of all the generators plus all possible combinations of them multiplied together. The aliases of any effect can be found by simply multiplying that effect times every term in the defining relation (Box and Hunter, 1961a:320-321).

As an example, the alias of $\vec{1}$, the main effect of factor 1, is determined as follows:

$$\vec{1} \quad \vec{I} = \vec{1} \cdot \vec{1}\vec{2}\vec{3}\vec{4}$$

$$\vec{1} = \vec{1}^2\vec{2}\vec{3}\vec{4} = \vec{I}\vec{2}\vec{3}\vec{4} = \vec{2}\vec{3}\vec{4} \qquad (6)$$

thus, the main effect of factor 1 is confounded with the interaction between factors 2, 3, and 4. As an example of a defining relation with more than one term, consider the 2^{7-4} design shown in Table IV. The generating relations are

$$\vec{I} = \vec{1}\vec{2}\vec{4} = \vec{1}\vec{3}\vec{5} = \vec{2}\vec{3}\vec{6} = \vec{1}\vec{2}\vec{3}\vec{7} \qquad (7)$$

the defining relations are

$$\vec{I} = \vec{1}\vec{2}\vec{4} = \vec{1}\vec{3}\vec{5} = \vec{2}\vec{3}\vec{6} = \vec{1}\vec{2}\vec{3}\vec{7} = \vec{2}\vec{3}\vec{4}\vec{5} = \vec{1}\vec{3}\vec{4}\vec{6} = \vec{3}\vec{4}\vec{7}$$

$$= \vec{1}\vec{2}\vec{5}\vec{6} = \vec{2}\vec{5}\vec{7} = \vec{1}\vec{6}\vec{7} = \vec{4}\vec{5}\vec{6} = \vec{1}\vec{4}\vec{5}\vec{7} = \vec{2}\vec{4}\vec{6}\vec{7}$$

$$= \vec{3}\vec{5}\vec{6}\vec{7} = \vec{1}\vec{2}\vec{3}\vec{4}\vec{5}\vec{6}\vec{7} \qquad (8)$$

The aliases of $\vec{1}$ are found to be

TABLE IV

DESIGN MATRIX FOR A 2^{7-4} DESIGN
(Box and Hunter, 1961a:320)

1	2	3	4=12	5=13	6=23	7=123
-	-	-	+	+	+	-
+	-	-	-	-	+	+
-	+	-	-	+	-	+
+	+	-	+	-	-	-
-	-	+	+	-	-	+
+	-	+	-	+	-	-
-	+	+	-	-	+	-
+	+	+	+	+	+	+

$$\hat{I} = 24 = 35 = 1236 = 12345 = 346 = 1347 = 256$$
$$= 1257 = 67 = 1456 = 457 = 12467 = 13467$$
$$= 234567 \text{ (Box and Hunter, 1961a:320-321)} \qquad (9)$$

If, in the analysis of the data, a factor is deleted, every term or word in the defining relation containing that factor is deleted. (In practice, a factor is deleted if it has no effect on the response.) If factors are deleted until no words remain in the defining relation, the design becomes a full factorial design in the remaining factors. For example, if any 5 factors are dropped from the 2^{7-4} design above, the design becomes a full factorial in 2 factors. Table V shows factors 1, 2, 3, 5, and 7 deleted in

TABLE V

THE EFFECT OF DELETING FACTORS FROM A 2^{7-4} DESIGN

Factor Deleted	Remaining Words of Defining Relation
1	$\vec{I} = \vec{2}\vec{4} = \vec{3}\vec{5} = \vec{3}\vec{4}\vec{6} = \vec{2}\vec{5}\vec{6} = \vec{6}\vec{7} = \vec{4}\vec{5}\vec{7} = \vec{2}\vec{3}\vec{4}\vec{5}\vec{6}\vec{7}$
2	$\vec{I} = \vec{3}\vec{5} = \vec{3}\vec{4}\vec{6} = \vec{6}\vec{7} = \vec{4}\vec{5}\vec{7}$
3	$\vec{I} = \vec{6}\vec{7} = \vec{4}\vec{5}\vec{7}$
5	$\vec{I} = \vec{6}\vec{7}$
7	NONE

that order. The resulting design is a full factorial in factors 4 and 6.

The designs in Tables III and IV demonstrate another advantage to using two levels in a factorial design. By using +1 and -1 to represent high and low values of the levels, each column in the design matrix is orthogonal. That is, if any column is vector multiplied by the transpose of any other column, the product is zero. Orthogonality insures that the estimates of the effects are independent and results in minimum variance estimates of those effects for the number of observations taken (Kleijnen, 1975:319).

In 1961, Box and Hunter first developed three special types of 2^k factorial designs:

(i) Designs of Resolution III in which no main effect is confounded with any other main effect, but main effects are confounded with two-factor interactions and two-factor interactions with one another. The 2^{3-1} design is of Resolution III.
(ii) Designs of Resolution IV in which no main effect is confounded with any other main effect or two-factor

interaction, but where two-factor interactions are confounded with one another. The 2^{4-1} design is of Resolution IV.

(iii) Designs of Resolution V in which no main effect or two-factor interaction is confounded with any other main effect or two-factor interaction, but two-factor interactions are confounded with three-factor interactions. The 2^{5-1} design is of Resolution V.

In general, a design of Resolution R is one in which no p factor effect is confounded with any other effect containing less than R-p factors [Box and Hunter, 1961a:319].

The notation adopted for these designs adds a Roman numeral subscript to the standard fractional factorial notation, e.g., 2^{4-1} (Box and Hunter, 1961a:319).

These designs have some especially nice properties for screening a few important factors from a larger group. First, if certain effects are expected to be relatively large, the factors can easily be arranged so that those effects are not confounded with each other. If no preliminary estimate is available, the arrangement is arbitrary. In most instances an arbitrary arrangement will allow at least a tentative identification of the important factors. If greater accuracy is still desired, the factors can be rearranged to allow better estimates of the effects. For example, in the 2^{7-4} design above, the main effect of factor 1 and the interaction between factors 2 and 4 are confounded with each other. If these two effects are large, and accurate estimates of each are desired, the factors in the design could be rearranged by reversing the positions of factors 1 and 7. Individual estimates of $\vec{1}$ and $\vec{24}$ could then be obtained.

Multiple runs of the experiment can usually be made at a reasonable cost due to a second "nice" property, the efficiency of the designs. That is, they require a relatively small number of runs to estimate the effects with some accuracy. For k factors, when k is a power of 2, fractional factorial designs of Resolution III exist that require only k+1 runs. If k is a factor of 4, but not necessarily a power of 2, Resolution III designs exist that are not fractional factorials, but still only require k+1 runs (Box and Hunter, 1961a:319-320). If k is not a power of 2 or a factor of 4, a Resolution III design may still be used by omitting variables from the Resolution III design of the next higher order (Box and Hunter, 1961a:323). Resolution IV designs require 2k runs. Orthogonal designs require k to be at least a factor of 4, but non-orthogonal designs are available which require k to be only a factor of 2. However, if a non-orthogonal design is used, the effects are mutually dependent and the variances are greater than they would be for an orthogonal design. In addition, the estimation of the effects is more complicated (Kleijnen, 1975:344-348). Resolution V design can include up to a certain number of factors for a given number of runs according to Table VI. However, in exchange for not confounding two-factor interactions with other two-factor interactions, Resolution V designs require so many more runs that they become relatively inefficient (Kleijnen, 1975:350). As a compromise between unconfounded effects and number of runs

TABLE VI

RESOLUTION V DESIGNS
(Box and Hunter, 1961b:149)

Number of Runs	Maximum Number of Factors
16	5
32	6
64	8
128	11

required, Resolution IV designs appear to hold the most promise for a moderate number of factors.

A third "nice" property of fractional factorial designs of Resolution R is that they provide a full factorial design in any R-1 factors. As discussed earlier, fractional factorial designs degenerate into full factorials when enough factors are deleted. For a Resolution R design with k factors, deleting k-R+1 factors always results in a full factorial in R-1 factors (Kleijnen, 1975:373).

Finally, fractions of full factorials are not unique. Alternative fractions of a family are created by reversing the sign of one or more of the generating relations. The fractions that result each allow the estimation of somewhat different combinations of effects. All of the fractions in a family taken together form a complete factorial (Box and Hunter, 1961a:322-323). This property allows two things. First, if the estimates provided by, for example, a 1/16 fraction are not accurate enough, another member of the family can be used to augment the first fraction. Second, it provides an alternative to rearranging the factors when unwanted combinations of effects are confounded. This technique would require considerably more work than rearranging the factors, but might be necessary in some instances.

Detailed descriptions of fractional factorial designs can be found in many books on experimental design. For this research, Box and Hunter's original work (1961a and 1961b) and Kleijnen's text (1975) provided clear and adequate expositions of the technique. In addition, Kleijnen provides extensive bibliographies on response surface methodology and the design and analysis of experiments. These include publications as recent as 1972.

Rotatable and Response Surface Designs. Rotatable and response surface designs were both developed to provide second-order or greater polynomial representations of response surfaces. Although Hunter and Naylor consider them techniques for general exploration, they provide enough detail about the response surface to be considered optimal-search techniques. Shannon and Kleijnen both consider them Response Surface Methodology (RSM) techniques (Shannon, 1975:1969-176; Kleijnen, 1975:356). As has been the case with every design considered, these designs achieve their increased accuracy by requiring a relatively large number of design points. In a rotatable design,

design points are used that are all the same distance from a central point. As a result, rotatable designs have some desirable properties relating to their prediction intervals. Rotatable designs require fewer design points than full factorials, but more than fractional factorials. Response surface techniques use a full factorial design on a subset of the factors, holding the remaining factors constant. If more than a few variables must be varied, a fractional factorial or rotatable design would have to be used instead (Hunter and Naylor, 1969:52). These designs might hold some possibility for use in validation once the important factors are found, but they require too many runs to be used in screening.

Screening Designs. In addition to fractional factorial designs of two levels, Kleijnen discusses three other types of screening designs. These designs are applicable in situations where, for example, k' important factors are to be screened from k conceivably important factors, where k' is much less than k, and k is very large, say 100 or more (Kleijnen, 1975:374). These designs are not applicable to the Georgia Tech model, but they might be used in screening for the important factors in TAC ZINGER. A brief description each will only be given here. For further discussion see Kleijnen (Kleijnen, 1975:374-407).

The first type are random designs, where all or some of the elements of the design matrix are chosen by a random sampling process. Levels of a particular factor are usually assigned equal probabilities of selection. However, if prior knowledge about certain levels indicates they are more promising, a non-uniform sampling distribution may be used. There is no connection between the number of observations and the number of factors, although the larger the ratio of the two, the better the estimates obtained. There are no specific techniques for analysis of these designs. ANOVA may be used on a restricted subset of the variables. Simple regression analysis of main effects only is possible by making the response a stochastic variable. Graphical presentations of the main effect of each variable on the response can provide a very good intuitive feel for which factors are important. Several significant tests may also be applied to supplement the graphical analysis technique. Kleijnen concludes that the major advantage of random designs is the small number of runs required. There are many disadvantages, which prompted the development of non-random designs that also require very few observations for a large number of factors (Kleijnen, 1975:376-386).

The next type of screening designs are supersaturated designs, which are incomplete fractional factorial designs in which the number of runs is less than the number of factors. As a result, such designs are not orthogonal, but are constructed to make them as close to being orthogonal as possible. Designs for 36 factors or less have been tabulated. An iterative algorithm to determine a design for more than 36 factors may be implemented on a computer, but the time required to do so may be prohibitive. In that case a random design may have to be resorted to (Kleijnen, 1975:386-393).

The final type are group-screening designs. Each group is then divided into several groups, which will each then be treated as a single factor. If necessary, multiple-stage procedures can be used with these designs. That is, groups that have a significant effect on the response can be repartitioned into smaller groups yet, and then tested again. Group-screening designs are the only designs that are recommended when the number of variables is very, very large. Kleijnen suggests that they might be used even when the number of variables is on the order of 100,000 or more (Kleijnen, 1975:393-407).

Kleijnen discusses one possible approach to screening that combines these techniques. When nothing at all is known about the response surface, he recommends randomly sampling the factors "to get some impression about the form of the response function, about the experimental region of interest and about the important factors." If this is not done, some factors must be assumed to have no effect or else held constant. Once the approximate region of interest is determined, a systematic orthogonal supersaturated, or group screening design can be used (Kleijnen, 1975:392-393).

For the Georgia Tech Model, 16 factors were identified as possibly being important. Their selection is justified in the next chapter. For 16 factors, a 2^{16-11}_{IV} should provide an appropriate balance between the number of runs required and the accuracy of results desired. As mentioned earlier, regression analysis will be used to analyze the data. A discussion of regression analysis as it applies to fractional factorial designs follows.

<u>Regression Analysis of Fractional Factorial Experiments</u>. The applicability of regression analysis to factorial designs is well-documented. Kleijnen's development will be followed here. The basis for using the approach is that "the [mathematical] models for factorial designs are special cases of the general linear regression model" (Kleijnen, 1975:299-301). The form of the model is

$$\vec{Y} = \vec{X}\vec{\beta} + \vec{\varepsilon} \tag{11}$$

For linear regression, $\vec{Y}$ is the vector of dependent variables or responses; $\vec{X}$ is the matrix of independent variables; $\vec{\beta}$ is the vector of least-squares estimators; and $\vec{\varepsilon}$ is a vector of independent, normally distributed random variables with zero mean and constant variance σ^2. For factorial designs, $\vec{Y}$ is the vector of responses; $\hat{X}$ is still the matrix of independent variables, but each element takes on a value of zero or one (plus one or minus one in the special case of only two levels per factor discussed later); $\vec{\beta}$ is the vector of estimates of the grand mean, the main effects, and the interactions; and $\hat{\varepsilon}$ remains the same as in regression (Kleijnen, 1975:301).

To illustrate, consider two factors A and B each with two levels, 1 and 2. Adopting Kleijnen's notation, the grand mean is denoted μ; the main effect of factor A at level i is α_i^A; for factor B, α_i^B; the two factors with A at level i and B at level j is α_{ij}^{AB} (Kleijnen, 1975:291-292). The formulation of the model is then

$$
\begin{aligned}
y_{11} &= \mu + \alpha_1^A + \alpha_1^B + \alpha_{11}^{AB} + \varepsilon_{11} \\
y_{12} &= \mu + \alpha_1^A + \alpha_2^B + \alpha_{12}^{AB} + \varepsilon_{12} \\
y_{21} &= \mu + \alpha_2^A + \alpha_1^B + \alpha_{21}^{AB} + \varepsilon_{21} \\
y_{22} &= \mu + \alpha_2^A + \alpha_2^B + \alpha_{22}^{AB} + \varepsilon_{22}
\end{aligned}
\tag{12}
$$

The results of this formulation may be analyzed using the following form of the regression equation:

$$y_{ij} = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2j} + \beta_3 x_{1i} x_{2j} + \varepsilon_{ij} \tag{13}$$

where y_{ij} = response with factor A at level i and factor B at level j;

x_{1i} = the value of factor A at level i;

x_{2j} = the value of factor B at level j;

β_0 = μ;

β_1 = α_1^A;

β_2 = α_j^B;

β_3 = α_{ij}^{AR};

and the error terms remain the same (McNichols, 1979:4-36).

The above formulation holds for full factorial designs, where there is no confounding of effects. In fractional factorials, effects are confounded and the formulation must account for this. In discussing confounding of effects earlier, it was noted that when two or more effects are confounded with each other, their individual estimates will all be equal. Therefore, if one affect is estimated, the values of its aliases are known, and in fact cannot be estimated separately. This is handled in the regression equation by simply including only one alias out of every set of confounded effects. For example, if the main effect of factor A and the two-factor interaction between factors B and C are confounded, either one may be included in the equation.

In the formulation of the general linear model, the assumptions are made that the experimental error is normally and independently distributed with zero mean and constant variance. Making these assumptions allows the F-test to be used in analyzing the data. What then, is the effect on the results if these assumptions do not hold? Kleijnen discusses these problems and cites several empirical investigations in arriving at the following conclusions. In simulation experiments, independence is not an issue, since independence of the errors can be assured by using different random number streams for each run. Normality and equality of variance cannot be assumed, but the F-test is not very sensitive to the lack of either or both (Kleijnen, 1975:303-304). The degree to which the Georgia Tech model adheres to these assumptions is discussed in a later chapter.

Interpretation of Analysis Results. Using a two-level factorial design to screen for important variables allows a straightforward interpretation of the results of the analysis. In most applications the least-squares estimators, the $\hat{\beta}$'s, cannot be used to directly measure the relative importance of the independent variables. They can, however, be used for this purpose when the levels of the factors in a two-level factorial design are set at the extremes of their expected ranges.

In regression analysis, the $\hat{\beta}$'s estimate the change in the dependent variable due to a one unit change in the respective independent variables. When the regression analysis is performed on the raw data as is most often done, the relative importance of the independent variables cannot be determined from the $\hat{\beta}$'s alone. This is due to differences in the scales and ranges of the independent variables. Consider, for example, a regression with only two independent variables, $Y = \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2$. Suppose X_1 ranges between .001 and .008, and X_2 between 100 and 1,000. Let $\beta_1 = 100$ and $\beta_2 = 10$. Running X_1 over its entire range from .001 to .008 would cause a change in Y of only (.007)(100) = .7. Running X_2 over its entire range would cause a change in y of (900)(10) = 9000. Clearly, if X_1 and X_2 can be expected to vary over the range considered, X_2 has a much greater effect on y than x_1, and $\hat{\beta}$'s do not reflect the relative importance of the variables.

In economic applications, the common way of solving this problem is to compare the elasticities of the variables. The elasticity is the percent change in the dependent variable divided by the percent change in the independent variable. In contrast to the $\hat{\beta}$'s, elasticity, provides a direct measure of relative importance of the variables. For example, elasticity is especially useful in analyzing economic data, since little or no control can be exercised over the values of the observed independent variables. In such applications, elasticity provides a method of standardizing the changes in the variables so their effects may be directly compared.

In computer simulations, the independent variables are controllable, and setting them at their high and low extremes allows a different way of standardizing those changes. Instead of using raw data for the regression analysis, the independent variables in the equation may be represented by +1 and -1. Thus, when the $\hat{\beta}$'s are calculated, a one unit change in the independent variables represents a change of half their reasonable range. Thus, the $\hat{\beta}$'s represent one-half the maximum effect the independent variables could have on the dependent variables. The relative importance of the independent variables can therefore be inferred directly from the $\hat{\beta}$'s.

The Multiple Response Problem. One problem that did not arise in this research, but can be serious in other applications, is the *multiple response* problem. That is, how is the analysis of data handled when there is more than one response variable of interest? Naylor, Burdick and Sasser phrase the situation quite succinctly.

> It is often possible to bypass the multiple response problem by treating an experiment with many experiments each with a single response, or several responses could be combined (e.g., by addition) and treated as a single response. However, it is not always possible to bypass the multiple response problem; often multiple responses are inherent in the situation under study. Unfortunately, experimental design techniques are virtually nonexistent [Naylor, et al., 1969:30].

Kleijnen (1975:408) also notes the lack of experimental designs that can handle multiple responses, and cites a few exceptions that have limited application. Shannon (1975:229-231) advocates the use of tolerance regions. No examples of their use were found. Naylor, et al. (1969:30) felt that utility theory holds the key to the solution to the problem. This concept would fit in well with Tytula's use of decision analysis in validation described in the previous chapter, and deserves further attention in the literature.

An Approach to Using Experimental Designs in Validation for OT&E

In screening the Georgia Tech model, a fractional factorial design should provide an adequate exploration of the response surface with the fewest design points. The specific design used will be discussed later. Provided that the screening process does show that only a few factors are important. The response surface for those factors can be defined in more detail, using one of the other designs discussed. The important factors can be varied over appropriate ranges, the unimportant factors can be held constant at some neutral (i.e., not extreme) value.

The next step in the validation process would be to run experiments with the actual system, concentrating on the important variables already identified from the simulation. As in the simulation, the unimportant factors should be set at a neutral level, and the important ones varied according to a valid design. Insight into the appropriate levels to use should come from the detailed map of the response surface generated by the simulation. These data points could then be compared against the simulation's predictions.

Confidence in the agreement between the simulation and the real system's data will hinge on the confidence in accuracy of the data compared. The simulation can be replicated using variance reduction (Monte Carlo) techniques to establish some level of confidence in the mean response for a given design point. However, the problem is not so easily remedied for the actual system.

First of all, each run of an experiment with the real system is expensive. Typically, a fixed amount of money is available to make those runs. Thus, a trade-off between data points and replications per data point must be made. On one hand, a sufficient number of design points are needed to adequately explore the response of the system under a variety of circumstances. On the other hand, design points with high variability in their response require replication to obtain a confident estimate of the mean response.

The data from the detailed description of the simulation's response surface can be of value here in determining the appropriate mix. At data points where the simulation's response exhibited high variability, the actual system's response can be expected to behave similarly. The same should hold for points of low variability. Thus, the simulated data can be used to estimate the replications required to establish a given level of confidence in the response for a particular design point. Alternatively, the simulated data can be used to determine the confidence level for a given number of replications.

Suppose the number of design points required to adequately describe the surface has already been determined, and cost constraints will not allow the desired number of replications to be made. The simulated data can be used to allocate available replications to those design points with the greatest variability. It can also be used to estimate the confidence level at those design points which are replicated fewer times than desired. If the variability is still higher than desired, a decision will have to be made: would fewer design points and more replications, or would the given design points with their individual accuracy result in greater confidence in the accuracy of the description of the system's response surface? Some form of decision analysis could again be employed in making that decision. Such an application is left for future research.

The second problem in increasing confidence in the accuracy of the actual system's results is that, especially for missile systems, it is extremely difficult to precisely control all of the input variables. Factors in a static system, such as frequency, beamwidth, and pulsewidth in a radar system, can be precisely controlled. However, factors in dynamic systems, such as exact altitude, velocity, or pitch or yaw angles, cannot be controlled. Additionally, it is impossible to duplicate environmental conditions unless the runs of an experiment are all made virtually simultaneously (Tytula, 1978:22). All of these factors add noise to the data, as does sheer error in collecting the data. The simulation cannot account for this noise other than by randomly adding it to the system. However, it will have already identified the factors that have major effects on the response, extra care can then be exercised to insure those factors are as closely controlled as possible. Application of such considerations falls outside the boundaries of this research.

One serious problem remains to be addressed in implementing the above approach. Experimental designs specify parameter combinations for very dissimilar circumstances. The Georgia Tech model can easily handle discontinuous changes in parameter settings from run to run; the actual system cannot. For the simulation, the parameter values specified for each design point are simply entered as input data. The simulation then generates a response for that particular set of parameters. In contrast, a missile must be flown under strictly controlled conditions in order to measure the actual system's response at each design point. Strictly following experimental design techniques discussed would make use of data only from those points on the missile's flight path that are specified by the design. The system's conditions specified for those design points may vary enough from one point to another to require a separate flight for each point. However, even if data for several design points can be gathered from one flight, to use only that data and ignore the rest of the flight is out of the question. Efficient use of resources requires obtaining as much usable data as possible from any run of a system.

Data from other than specified design points can obviously be used to test the agreement between the simulation and the actual system. But if it is, the same problems arise as when no experimental design is used. First, at which of those added points should the simulation be run to compare the data? Second, once those points have been determined, the agreement will not necessarily hold under other conditions. Unlike the case of no experimental design, however, these test runs have not been arbitrarily chosen. They have been designed to obtain data at points that are representative of the full range of the system parameters. Comparisons of simulated and actual data can be made at points from a cross-section of these runs. Comparisons at these points should be a better test of the ability of the model to predict the actual response in situations not tested than if the points were from flight paths determined without consideration of any experimental design.

This approach, too, neglects a great deal of information available from each run. This is because the number of discrete points that can be compared is limited by cost constraints. The solution to this problem most likely lies in some technique that compares entire flight paths. Such a solution would probably require expressing both the simulated and actual data as time series. Provided the data can be expressed in time-series form, it can be analyzed using standard techniques. However, for OT&E applications, Tytula's approach, described in the previous chapter, should be considered before attempting to apply other techniques.

The approach described in the following chapter was directed toward identifying the important factors of the Georgia Tech model using a fractional factorial design.

IV. Research Methodology

The purpose of this research was to develop insight into methodology that could be used in validating TAC ZINGER and TAC REPELLER. The research done was based on two hypotheses: first, that out of the many variables in the models, only a few were really important; and second, that if the research were done with a submodel of TAC ZINGER, that research would provide some insight into validating the larger models. Both of these hypotheses were necessitated by the size and complexity of the larger models.

To test the first hypothesis, experimental design techniques were used to systematically identify the important variables and statistically test their significance. The degree to which the second hypothesis holds cannot be fully determined until the larger models are validated. However, information gained and lessons learned from this research that have potential applicability to the larger models will be discussed.

The Georgia Tech Model was chosen as the submodel to be used for several reasons. It was of reasonable size and complexity for the scope of the research. It was currently being used in an operational program, TAC ZINGER, but could be extracted and run independently. The computer code for the program was independently documented by a current report (Stuk, 1979a). The model had already undergone some validation testing (Zehner and Tuley, 1979). These considerations led to the assumption that the model's independent implementation would be relatively easy and require minimal verification of its proper operation. In reality, verification of the program required the majority of the time available for the research.

The research with the model itself had five objectives. The first was to properly implement the program independent of TAC ZINGER. The second, to identify from the documentation those parameters that would affect the model's output and could be controlled. The third, to determine the appropriate experimental design to use with those variables. These latter two could be done while the first was being done. The fourth, to generate data with the model, using the design previously determined. The final objective was to analyze the results using regression analysis.

The Georgia Tech Model has thus far been described only very generally. To follow the discussion further, a somewhat more detailed description is needed, and follows.

Model Description

The set of subroutines that have been called the Georgia Tech Model in this thesis are used in TAC ZINGER to calculate the effects of clutter and multipath on radar tracking error. The subroutines calculate those effects based on characteristics of three elements of the

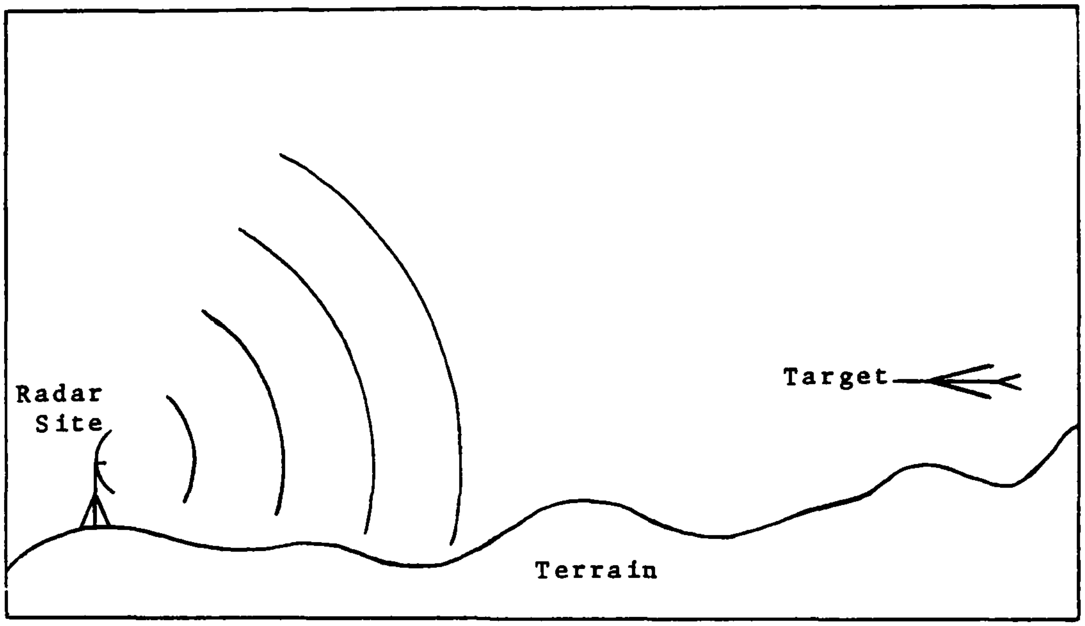

Figure 3. Model Scenario

surface-to-air missile engagement scenario: a radar tracking site, an airborne target being tracked by the radar, and the terrain between the two (Figure 3).

When the target is at a low altitude, clutter and multipath can seriously degrade the radar's tracking ability. Clutter is unwanted radar return that has been backscattered from the terrain surrounding the target. It "provides a competing signal to the target return," and is

> . . . dependent on terrain type, depression angle, surface roughness, and radar characteristics. . . . Multipath is a consequence of the multiple paths by which a signal may complete the round trip from radar to target and back to radar [Zehner and Tuley, 1979:1].

Figure 4 illustrates those paths.

<u>Model Scenario</u>. The characteristics of the radar site, terrain, and target are all specified prior to calling on the model to perform its calculations. The terrain

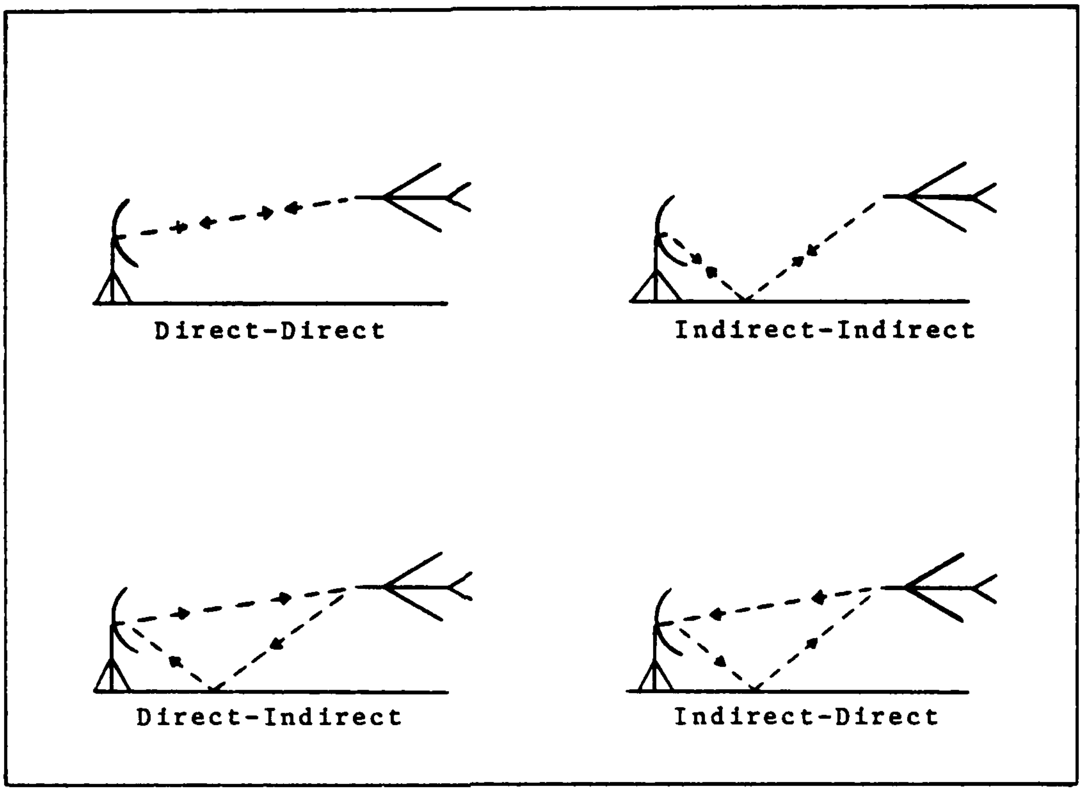

Figure 4. Possible Radar Signal Paths (Tuley, 1979a)

characteristics are read from an external data file. The target and site characteristics are specified by TAC ZINGER when the model is called. The model uses this information in the appropriate equations to generate the radar tracking error. The mathematical model for the low-angle radar tracking situation was developed by D. K. Barton (1974: 687-704).

Some of the radar's characteristics are passed to the model directly through the call from TAC ZINGER; the remainder are assigned by the model, based on a parameter passed in the call. The characteristics passed directly are the antenna's location and the direction it is pointing.

The location is specified as coordinates in feet in a three-dimensional reference coordinate system. The coordinate system is a standard right-hand X-Y-Z system. The direction the antenna is pointing is specified by its azimuth and elevation angles. The azimuth angle is measured counter-clockwise in the X-Y plane, beginning at the X-axis. The elevation angle is measured up from the horizontal. The model contains appropriate values for the frequency, pulse-width, and azimuth and elevation beamwidths of several types of radar systems. In the initial call to the model, TAC ZINGER passes a parameter, KEYSAM, which specifies the system to be used.

The target's characteristics are completely specified in the call to the model. Its location is defined in the same way as the site's, and the azimuth and elevation to the target from the antenna are specified. Because the radar does not track the target perfectly, these angles will not usually be the same as the site's elevation and azimuth angles. The target's physical characteristics are represented in the model only by its radar cross-section, which is also passed in the call.

The terrain is approximated in the model by a collection of flat plates called facets. Each facet is 1650 feet square, and is defined by its location in the X-Y plane, height above the X-Y plane at the facet's center, slopes in the X and Y directions, root mean square (RMS) surface roughness, and terrain type. Figure 5 shows a

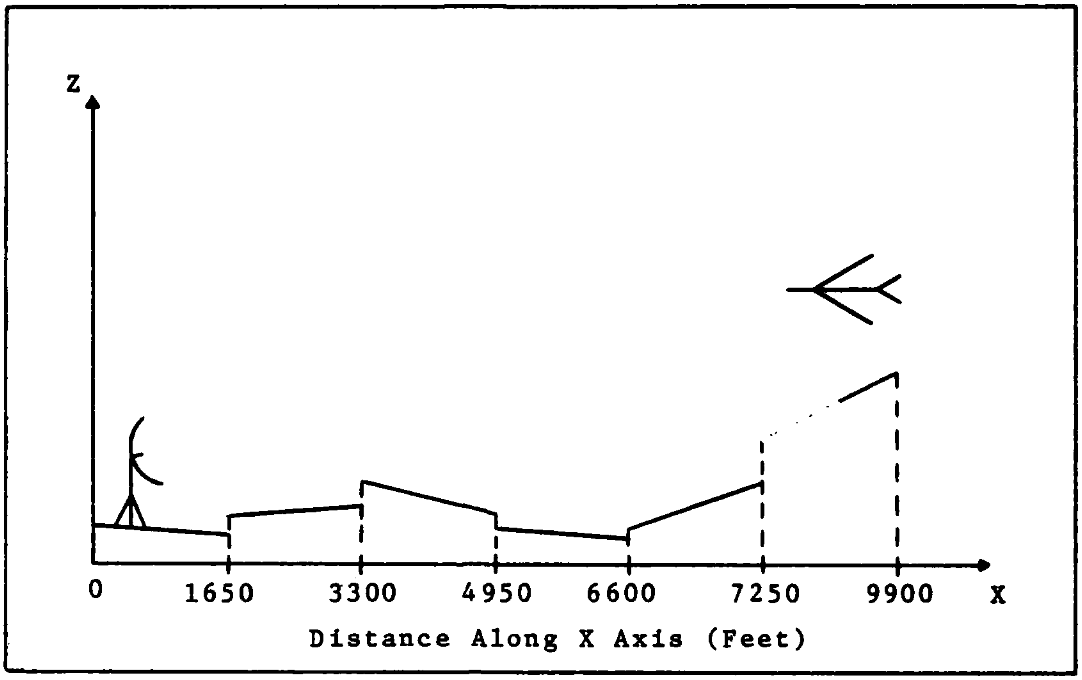

Figure 5. Cross-sectional View of Terrain as Approximated in the Model

cross-sectional view of terrain along the X-axis as it might be approximated in the model. Discontinuities at the facet edges occur as the result of using flat plates to approximate actual terrain.

To define a facet's slope, or tilt, consider a set of axes, X'- Y'- Z', parallel to the reference coordinate system, but whose origin is located at the center of the facet surface (Figure 6). Define a vector normal to the facet, $\hat{F}$, at that point, and decompose it into components F_{xz} and F_{yz} in the X'-Z' and Y'-Z' planes respectively. The slope in the Y direction, YTILT, is defined similarly. The sign of either angle is determined by further decomposing F_{xz} and F_{yz} into their components along the X' and Z'

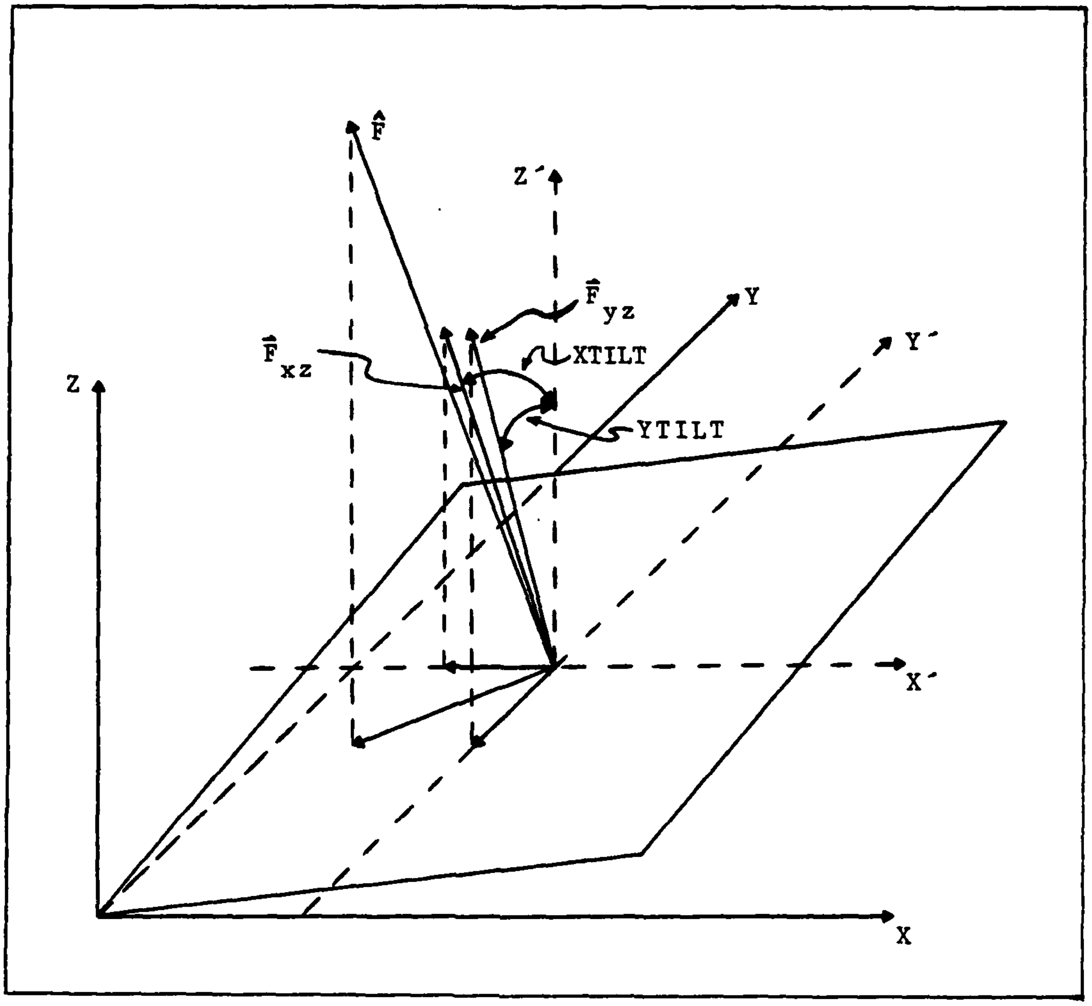

Figure 6. Facet Slope Geometry

and Y´ and Z´ axes. The sign of YTILT corresponds to the sign of F_x, and the sign of YTILT corresponds to the sign of F_y.

For actual terrain, the facet height, slope, and RMS surface roughness are determined by using a least squares technique to fit a plane to a number of measured elevations for each facet, e.g., 25. The facet's center height and slope are calculated directly from the equation of the plane.

The RMS surface roughness is the square root of the mean square elevation error between the plane and the actual terrain at the data points (Tuley, 1979a). Generating the data required to map actual terrain into a form usable by the model is a tedious and time-consuming process. To allow the model to be used when actual terrain data is not needed, the model provides a default terrain base. The default terrain is a 40-facet-square flat, rather smooth, surface. All facet heights and slopes are zero with an RMS roughness of 10 centimeters. The terrain type must be specified by the user.

The model allows the terrain to be specified as sea or one of eight land types, each of which may be wet or dry. Each land type is characterized by a set of empirically determined parameters that define its clutter characteristics, and by its dielectric constant, a measure of the degree to which it absorbs the radar energy that strikes it. The values for the land clutter parameters and dielectric constants are read from an external data file. The dielectric constant for sea is read from the data file along with those for land but its clutter characteristics depend on different parameters than land's and are included in the model itself.

A planview of the geometry of the scenario is shown in Figure 7. In calculating multipath, the model uses facets that are intersected by the ground line, the line that connects the X-Y planar coordinates of the site and

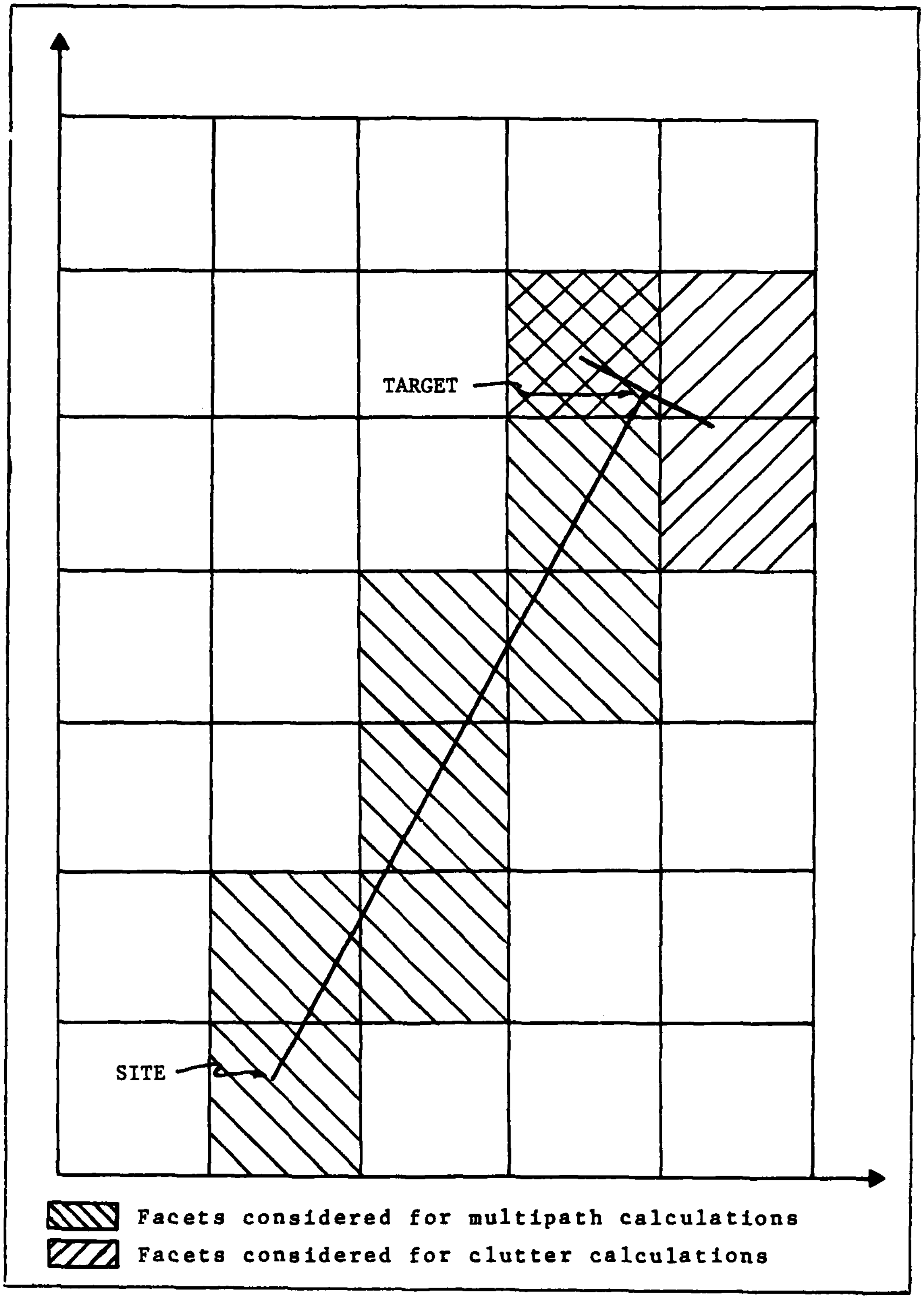

Figure 7. Planview of Scenario Geometry (Stuk, 1979a:12,14)

the target. The clutter calculations use facets intersected by a line segment perpendicular to the ground line at the target end. The length of the line segment equals the radar azimuth beamwidth in feet at the distance from the site.

Program Flow. As described briefly in Chapter I, TAC ZINGER simulates the engagement of an ingressing aircraft by a SAM site. As TAC ZINGER "flies" the target aircraft through the SAM environment, it has the site radar check at designated intervals of time to see if it has acquired the target. If it has, it tracks the target and launches a missile. The radar tracking calculations are performed by the Georgia Tech Model. The Georgia Tech Model returns to TAC ZINGER the elevation and azimuth track angle errors, and a target radar cross-section modified due to clutter and multipath effects. These values are used in calculating the outcome of the engagement at each time interval. At the next time interval, those values are used to modify the antenna's azimuth and elevation angles and the radar cross-section of the target. Thus, over the entire flight path, the output of both TAC ZINGER and the Georgia Tech Model are autocorrelated. However, at any given time interval, the Georgia Tech Model calculates its results based only on the parameters passed to it. It does not matter to the model that those parameters were or were not based on the results of a previous call. It is this property that allows the model to be run using the experimental design that is to be described shortly.

The basic flow of the model is shown in Figure 8. Subroutine MULTIN is called only one time to initialize parameters that will not change during the entire TAC ZINGER run; the terrain data is read, and the system-dependent characteristics of the radar are assigned. MULTIP is called at each time interval to generate the track angle errors and modified radar cross-section. In the call to MULTIP, TAC ZINGER specifies the three-dimensional locations of the site and target, the target's radar cross-section, the angular location of the target with respect to the site, and where the antenna is pointing. MULTIP calls FACET twice during each run. On the first call, FACET determines the facets capable of supporting multipath (Figure 7). MULTIP then calculates the reflection coefficients from those facets due to multipath. The second call to FACET determines the facets which could contribute to the clutter return. MULTIP uses information about those facets to calculate the radar cross-section of the clutter patch. MULIP calls four other subroutines in calculating other modifications to the multipath reflection coefficient and the radar cross-section of the clutter patch. Subroutine RICE generates random deviations in the multipath reflection coefficients and clutter patch radar cross-sections from a Rice distribution. (See Brookner (1977:391) for the definition of the Rice distribution.)

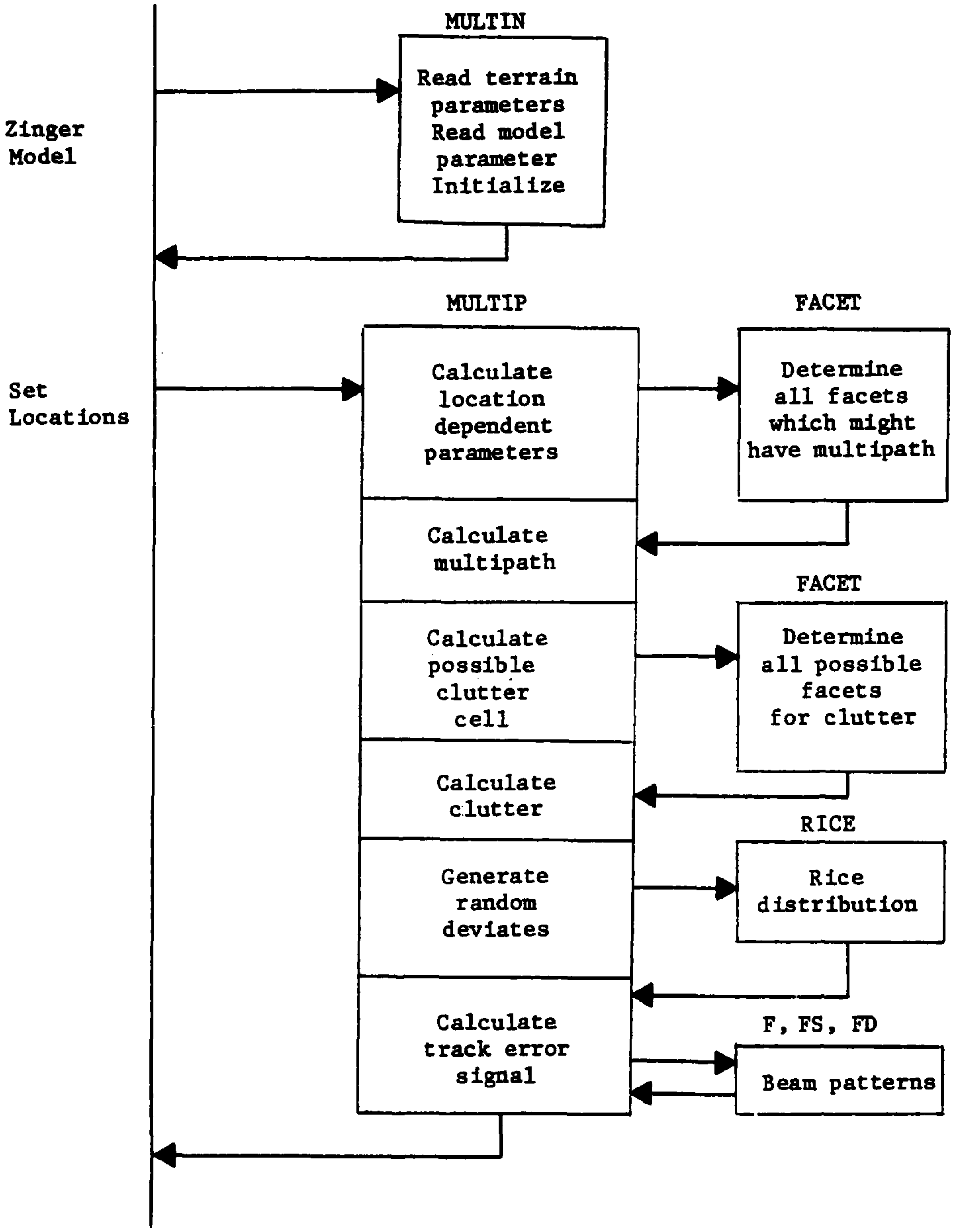

Figure 8. General Program Flow (Stuk, 1979a:3)

Subroutines F, FS, and FD are used to calculate the radar beam pattern and the sum and difference patterns, respectively. Finally, MULTIP uses this information to calculate the azimuth and elevation track angle errors and the modified radar cross-section of the target, which are returned to TAC ZINGER.

Model Documentation

The Georgia Tech Model is documented in two reports: one documents the computer programs (Stuk, 1979a); the other describes the development of the model and the validation testing it underwent (Zehner and Tuley, 1979). The theoretical development is well presented. As noted earlier, the model is based on work published by D. K. Barton in 1974. The research for this thesis assumed that the mathematical model was correct, and did not try to test that assumption. The software documentation, although much better than the documentation for TAC ZINGER, has some deficiencies. In addition, two programming errors were discovered in the model during the research. They are discussed later.

The computer programs for the model included in the software documentation have very few comments in them. This is to minimize the length of the programs for inclusion in TAC ZINGER (Stuk, 1979b). It would have been very helpful if the documented version had included appropriate comments, or if the documentation had included a more detailed description of the logic. The flow charts of each subroutine are very useful, but more detail is needed yet.

The model's documentation inadequately defines several key input variables. AZS and ELS are defined simply as the antenna azimuth and elevation in radians, AZT and ELT as the target's azimuth and elevation in radians (Stuk, 1979a:62). These have already been defined in a previous section of this chapter. The brief definitions given in the documentation are not elaborated on further, nor is there any mention of the point that TAC ZINGER modifies AZS and ELS based on the azimuth and elevation track angle errors calculated in the previous call to MULTIP. Furthermore, the reference point for defining AZT and ELT is not specified. When the site is at the origin of the reference system, there is no problem in defining AZT. But when the site is away from the origin, is AZT defined in relation to the site or the origin of the reference system? In either case, is ELT the elevation angle of the target measured from the antenna height, the terrain elevation, or the reference (X-Y) plane? The problems encountered as a result of these ambiguities are discussed in a subsequent section.

Another pair of parameters whose poor definitions resulted in several problems are XTILT and YTILT, the slopes of the facets in the X and Y directions. Again, these have already been defined here in a previous section. In the documentation, they are twice defined as simply the

AD-A083 516 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOO--ETC F/G 9/2
TOWARD VALIDATION OF COMPUTER SIMULATION MODELS IN OPERATIONAL --ETC(U)
DEC 79 J M ARNETT

UNCLASSIFIED AFIT/GOR/SM/79D-1 NL

2 OF 2
AD-
A083516

END
DATE
FILMED
5-80
DTIC

FACET tilts in milliradians in the X and Y directions (Stuk, 1979a:51). That reference is in the middle of a derivation of the coordinate system transformation used in calculating the angle errors. The convention for positive and negative angles is not mentioned in any of those locations. Assuming a wrong convention is easy to do and hard to detect. The program performs its calculations based on the terrain data it reads in, and does not know or care if that terrain represents reality or not. Two wrong conventions were assumed in the course of the research, and are discussed in a later section.

A very subtle, but extremely important discrepancy exists between the way the elements of the array ITARA, which stores the terrain characteristics, are defined and the way they are used in subroutine FACET. They are defined in the documentation as ITARA (K, I, L), where K is the terrain characteristic, I is the X index, and J is the Y index of the facet (Stuk, 1979a:58). In MULTIN, the terrain is loaded in ITARA (K, I, J), giving the impression that I, J corresponds to standard X, Y order. Finally, in FACET, ITILT takes on the value of the FACET's XTILT and JTILT takes on the YTILT value. The problem surfaces here. ITILT is assigned the XTILT value of ITARA (2, M, N), where <u>M is the Y index, and N is the X index</u>. <u>The indices are reversed from the expected X,Y order</u>. The same situation holds for the other four terrain characteristics. Thus, the terrain could be loaded in ITARA with the X and Y coordinates unknowingly reversed.

The reason for loading the terrain in Y,X instead of X,Y order has to do with the way FORTRAN searches an array. When a particular element in an array is called for, FORTRAN indexes the outer index first, then the inner. TAC ZINGER is generally run with the target aircraft flying parallel or close to parallel to the X axis. As a result, the facets considered in the multipath calculations have a wide range of X indices, but a relatively small range of Y indices. Given FORTRAN's search technique then, it is more efficient to store the terrain in Y,X order. Whatever the reason, such inconsistency in notation should have been highlighted in both the documentation and the computer program itself.

Another problem had to do with small-angle approximations. Since the model is for low-altitude scenarios, it is able to take advantage of small-angle approximations for the sine or tangent of many angles. At very small angles, the sine and tangent are almost equal to the angle itself (in radians). The computer program makes maximum usage of these approximations in order to reduce the program's execution time. However, the specific locations in the program where they are made are not always documented. This resulted in some initial confusion in the interpretation of parts of the code.

One line in MULTIN was cause for some concern that it did not include all the proper terms. The lines from the code including comments, are:

```
C      PULSEWIDTH IN NANO SECONDS
       C3=PW*.5*TS
C      C*TAU/2 IN TSCALE UNITS
```

where C3 is a constant to be used in MULTIP, PW and TAU both refer to pulsewidth, TS is 1/(1650 feet), and C is the speed of light. Until the dimensions on the variables are checked, it appears that the speed of light has been left out of the equation. The key to resolving the apparent problem is in recognizing that the speed of light is approximately one foot per nanosecond. In those units, the numerical value of the term "drops out" of the equation, but the dimensions remain. However, the dimensionality used is not documented in either the program or the documentation. An explicit statement of the units used would have eliminated the present potential for misinterpretation.

These last two points are relatively minor concerns as compared to the necessity of adequate variable definitions. However, they can, and in fact did, cause a lot of time to be lost in debugging the program. Understanding the logic behind an unfamiliar program is difficult enough without adding the confusion of undocumented assumptions and approximations.

In addition to the theoretical development of the model, the final report documents the validation work done on the model (Zehner and Tuley, 1979:63-93). The methodology employed was reasonable, but required making some assumptions that may not be valid, and which another methodology might not require. The model's results were compared against actual data from test runs with aircraft over sea and land (Zehner and Tuley, 1979:64). Only data on elevation error was collected; none on azimuth error (Tuley, 1979a). The procedure used divided the flight paths up into 500-meter segments and then aggregated the data from all of the runs into groups based on those segments. Within these groups the root mean square elevation error was calculated and plotted versus the median elevation angle of the targets within the group. A least squares fit equation of track error correction for range was made using data from reference runs at altitudes high enough to assume the multipath effect would be zero. The equation was assumed to hold for all the ranges and elevation angles considered. The equation was applied to the RMS elevation angle error for each 500-meter range group, so that the data could then be expressed as RMS elevation angle error versus elevation angle only. The final transformation was to divide the angles and angle errors by the beamwidth to facilitate "the extrapolation of information to radars of other beamwidths"(Zehner and Tuley, 1979:67-69). The same basic procedure was followed in transforming simulated data from runs that mimicked the actual test flights. Finally, the goodness of fit of the simulated to the actual data was evaluated using the F-test (Zehner and Tuley, 1979: 69-70).

This approach required several assumptions about the relationships between variables. First of all, it assumed that there was no interaction between range and any of the other variables. For example, in making the error correction for range in the way described, it was explicitly assumed that the correction for a given range was the same for all elevation angles. Second, in normalizing the data to beamwidths, it apparently assumed no interaction between beamwidth and any of the other variables, and that the main effect of beamwidth on angle error was a simple multiplicative relationship. Since the approach apparently did not address several factors, such as frequency, pulsewidth, surface roughness, and height of the antenna, it implicitly assumed their effects to be negligible. While these assumptions about variables and relationships may be valid, there was no way to test them with the methodology used.

As discussed earlier, input parameters in validation testing need to be chosen to explore the system's response over the entire expected ranges of those variables. The documentation did not discuss the reasons for choosing the test flight profiles used, although it appears that both range and elevation angle were run over a sufficiently wide range of values. Even so, application of experimental design techniques might have resulted in more efficient use of the test runs available.

This research did not make such assumptions about the relationships between variables. It used experimental design techniques to determine what those relationships are in the model. The variables considered and chosen for testing and their appropriate ranges are discussed below.

Variable Selection

The variables to be evaluated were determined from the model documentation, in conjunction with research to determine an appropriate experimental design. The variables that were candidates for testing have already been discussed as the characteristics of the radar system, the target, and the terrain. Reasons particular variables were not chosen are given in the discussion that follows. A 2^{k-p} fractional factorial design was chosen, and is discussed in detail in a following section.

The two levels chosen for each factor were the high and low extremes of its range. Reasonable values for the variables used were determined with the help of Major R. L. Broderson of the Cruise Missile Testing branch of Aeronautical Systems Division (ASD/ENFX), and Mr. M. T. Tuley of the Engineering Experiment Station, Georgia Institute of Technology. Major Broderson was familiar with the cruise missile's characteristics and with the characteristics of radar systems that might be encountered in an actual scenario. He was able to provide values for the target and radar parameters that were reasonable and yet unclassified. Mr. Tuley was familiar with work that had been done in approximating terrain for use in the model and so was able to provide reasonable extreme values for the terrain parameters. In addition, they were able to provide information on the expected response of the model over the ranges of most of the variables.

Response Variables. The independent variables selected had to be evaluated on the basis of their effect on some response. The three output variables, elevation and azimuth track angle errors and modified radar cross-section were obvious candidates. Although not so readily accessible, the multipath reflection coefficients and the radar cross-section of the clutter patch could also be used. Since in actual use of the model primary emphasis has been placed on elevation track angle error (Tuley, 1979b), it was chosen as the primary response against which the independent variables would be evaluated. For the information of future users, statistics on the other two responses are tabulated but not discussed.

Target Altitude. For experimental design purposes, the target's altitude was specified as height above the terrain. Since the model uses height above the X-Y plane, the appropriate transformation was made in the driver program before the model was called. Extreme values for the cruise missile scenario were chosen as 100 and 1000 feet. As altitude increased, the track error was expected to decrease (Broderson, 1979).

Target Range. The range of the target from the site was expected to be one of the most important factors chosen. As with target altitude, the model expresses the target range as coordinates in the reference system. The appropriate transformation is made prior to calling the model. Fifty thousand feet was used as the upper range. Since the model was developed for angles less than 20°, and due to the small angle approximations in the program (Stuk, 1979a:2), the maximum elevation angle of the target had to be restricted to 20°. Range and altitude are restricted by

$$\frac{\text{Altitude}}{\text{range}} \leq \tan 20°$$

At low ranges, the 1000-foot altitude already chosen is the critical value. At a target altitude of 1000 feet, the range must then be at least 3748 feet. Three thousand feet was chosen to satisfy the limitation.

The ranges chosen are reasonable for the cruise missile/SAM engagement scenario. Fifty thousand feet is just over 8 nautical miles, and 3000 feet is just under one-half nautical mile. Eight nautical miles is a reasonable outer range to expect a tracking radar to acquire and track a target as small and low as a cruise missile. Half a nautical mile is a reasonable inner range for the radar to be able to track the target and launch a SAM before the range starts increasing as the target passes the site. This is true even if the target were to fly directly over the site, a worst case and least probable situation (Broderson, 1979).

Although elevation angle was used as an independent variable in the documented validation effort, it was actually generated from range and altitude measurements. Elevation angle was thus not used, so that the effects of range and altitude could be estimated separately.

The expected behavior of the track error with range was not strictly monotonic, although the error was expected to generally increase with increasing range. Due to multipath effects, especially for a smooth terrain surface, elevation angle or error is expected to oscillate as it increases with range, similar to Figure 9 (Tuley, 1979b; Broderson, 1979). The absence of this expected behavior in results obtained in this research resulted in extensive debugging procedures to determine the cause. Late in the research, a change was implemented that corrected the problem. That change is discussed later.

Radar Cross-section of the Target. Reasonable values were chosen as .01 and 10 square meters. Ten square meters is actually a larger cross-section than a cruise missile would have, but was recommended as an upper level by Major Broderson (Broderson, 1979). The model accepts the radar cross-section in square feet expressed in decibels (dBsf). The transformation is made prior to reading the value into the driver program. The transformation from

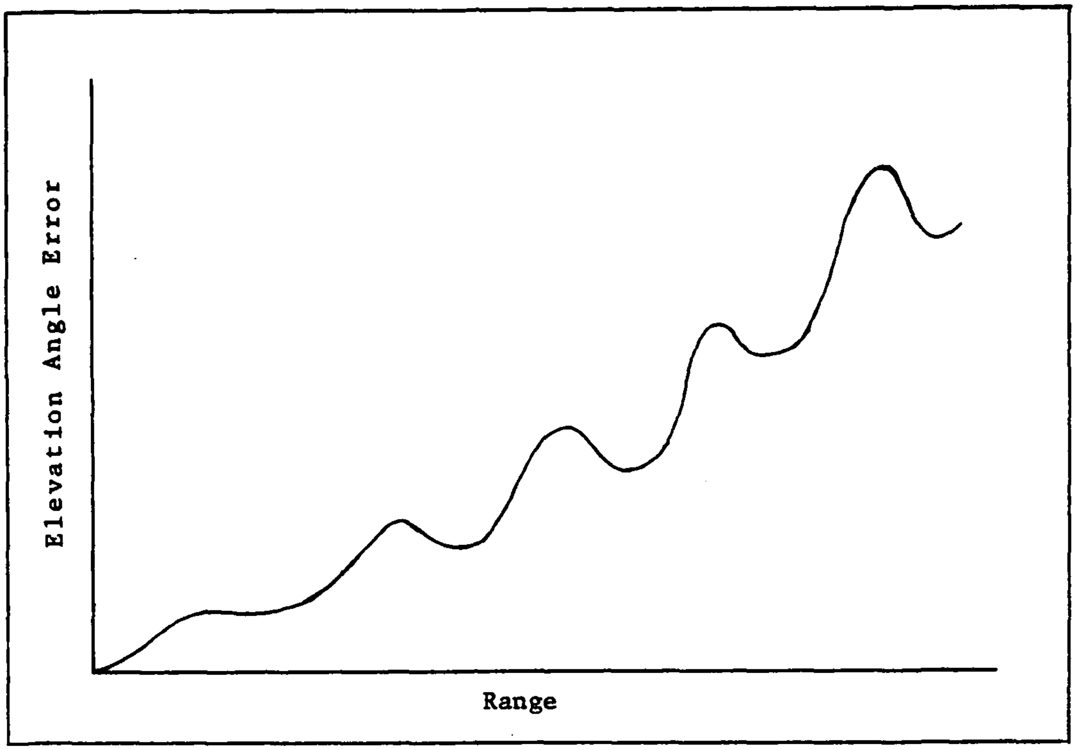

Figure 9. Expected Behavior of Elevation Angle Error with Range

square meters to dBsf is

$$\text{dBsf} = 10 \log_{10} \left[(\text{square meters}) \left(\frac{3.28 \text{ feet}}{\text{meters}} \right)^2 \right]$$

The track angle error was expected to decrease as the radar cross-section increased.

<u>Radar Frequency</u>. The software documentation states that "the model was developed for a Ku-band monopulse radar" (Stuk, 1979a:2). However, provisions are made for making clutter calculations at other frequencies (Stuk, 1979a:68). Additionally, the model is commonly used with frequencies from other bands (Table VII) (Broderson, 1979). The model, then, although developed for Ku-band, should be tested

TABLE VII

RADAR BAND DESIGNATIONS (Long, 1975:2)

Band	Frequency	Wavelength
P	300 - 1000 MHZ	30 - 100 CM
L	1000 - 2000 MHZ	15 - 30 CM
S	2000 - 4000 MHZ	7.5 - 15 CM
C	4 - 8 GHZ	3.75 - 7.5 CM
X	8 - 12.5 GHZ	2.4 - 3.75 CM
Ku	12.5 - 18 GHZ	1.67 - 2.4 CM
K	18 - 26.5 GHZ	1.0 - 1.67 CM
Ka	26.5 - 40.0 GHZ	.75 - 1.1 CM

over the range which it is actually used. Frequencies of 9 gigahertz (GHZ) and 16 GHZ were chosen to accomplish this (Broderson, 1979). The effect of frequency on angle error was not known.

<u>Radar Beamwidth</u>. Provision is made in the model for both elevation and azimuth beamwidths. It is not unreasonable to assume they are equal (Broderson, 1979), and that assumption was made for this research. Values chosen were .8° and 1.2° (3 dB beamwidth)(Broderson, 1979). The angle error was expected to increase as beamwidth increased.

<u>Radar Pulsewidth</u>. The pulsewidth was set at 100 and 300 nanoseconds (Broderson, 1979). As pulsewidth increased, angle error was expected to increase.

Sub-Clutter Visibility. Sub-clutter visibility is not defined in the body of the report. In the list of the input parameters its dimensions are given as decibels, but nothing more is said about it (Stuk, 1979a:62). Skolnik (1962:140) defines it as a measure of performance of moving-target-indication (MTI) radar. It is the gain in the signal-to-clutter power ratio produced by the MTI, measured in decibels (dB). For example, a radar that has a sub-clutter visibility of 30 dB can detect a moving target in the presence of clutter even though the clutter echo power is 1000 times the target echo power.

Appropriate values were chosen as 1 and 50 dB. As sub-clutter visibility increased, angle error was expected to decrease (Broderson, 1979).

Angle Off Boresight. The angle off boresight of the target is the difference between where the antenna is pointing and the actual direction to the target. Due to a misunderstanding that was not corrected until it was too late to do anything about it, angle off boresight was not considered as a parameter that could be varied. The misunderstanding was in the definitions of AES, ELS, AZT, and ELT in the call to MULTIP, as discussed earlier. Future research should include these parameters in the experimental design.

Georgia Tech commonly uses angle off boresight equal to zero in its runs (Stuk, 1979b), so that value was used for this research.

Terrain Type. Each of the nine different types of terrain is modeled by a specific set of clutter parameters and dielectric constants. Some of the clutter parameters are dependent on radar frequency as well as terrain type. In order to measure the effect of the terrain type separately from the effect of the radar frequency, the clutter parameters and dielectric constants were directly used in the experimental design. That is, values for the parameters themselves were directly specified instead of simply specifying the terrain type. This could be done because the code does not require the clutter parameters for each terrain type be kept together as a group. The values could thus be specified independent of each other, even though the combinations used did not correspond to actual terrain.

In the program, the radar cross-section of the clutter patch was first calculated for dry terrain. For wet terrain, the radar cross-section of the clutter patch is calculated by simply multiplying the dry cross-section by $10^{.5}$ (Zehner and Tuley, 1979:29). Since the results from wet terrain are thus dependent on those from dry terrain, only dry terrain values were used for the clutter parameters.

The values chosen for each parameter were its high and low values across all the terrain types. Table VIII lists the values chosen and the terrain types each value represented. Sea was not evaluated because the calculations of the radar cross-section of the sea clutter patch do not

TABLE VIII

CLUTTER PARAMETERS AND TERRAIN TYPES

Clutter Parameter	Low Value/Terrain Type	High Value/Terrain Type
A	.0097/9	2.0/8
B	.83/2,6,7	1.8/8
C	.0013/2,6,7	.015/8
D	0.0/3,4,5,8,9	2.3/2,6,7

Terrain Types

2 = Soil
3 = Grass
4 = Crops
5 = Trees
6 = Sand
7 = Rocks
8 = Urban Areas
9 = Wet Snow

use the clutter parameters in Table VIII. Future research should evaluate sea versus land results. Track angle error was expected to increase as the values of the clutter parameters increased.

Restricting the clutter calculations to strictly dry terrain does not necessarily restrict the multipath to dry terrain. The dielectric constants used in the multipath calculations have dry and wet values that are independent of each other. Thus, the results of the multipath calculations for wet terrain are not dependent on the results for dry terrain. The values for dry terrain are all fairly close together, as are the values for wet. Therefore, a low value of 2.0, corresponding to dry grass, was chosen. The high value was 20.0 + 2.4i, corresponding to wet soil. Using these values allowed an evaluation of the effect of multipath for dry terrain versus the effect of multipath for wet terrain. Track angle error was expected to decrease as the value of the dielectric constant increased.

A second run was made to evaluate the effect of the dielectric constant on dry terrain. The low value was the same as before. The high value was 2.55 + .016i, corresponding to sand.

Terrain Slope. The normal input to the model specifies the terrain slope in the X and Y directions, as defined earlier. Before the slope is used in any calculations, though, it is transformed into longitudinal and transverse tilts. The longitudinal axis is defined along the line from the site to the target coordinates in the X-Y plane. The transverse axis is perpendicular to the longitudinal axis. (The transformation is accomplished by rotating the X and Y axes through some angle, B, so that the X axis is aligned along the line between the site and the target. Appendix A gives a geometric derivation of the transformation.)

Since the program bases its calculations on longitudinal and transverse slopes, they were chosen for use in the experimental design. To match the model's inputs, though, they were transformed into X and Y tilts in the driver program. For both longitudinal and transverse slopes, extreme values of 0 and -10 degrees were used. (Due to the convention used, -20° slope indicates that the terrain rises as the distance from the origin increases in the first quadrant.) Zero degrees was used as the upper bound to insure the surfaces of its clutter facets could be seen from the site. Minus twenty degrees was a reasonable lower limit due to the facet size and method used to approximate the terrain (Broderson, 1979; Tuley, 1979b).

As longitudinal slope increased, the clutter return was expected to increase, but the multipath was expected to decrease. The relative effects of those changes was unknown, and so the expected change in angle error was not readily determinable. As transverse slope increased, both clutter and multipath returns were expected to decrease, and so the angle error was expected to decrease also.

Surface Roughness. Based on previous work done in mapping actual terrain into usable data for the model, extreme values for RMS surface roughness were chosen as 10 and 500 centimeters (Tuley, 1979b). Increased surface roughness was expected to increase the clutter return. The effect on multipath was indeterminate, since increased surface roughness was expected to have opposite effects on the two components of the multipath return. Specular multipath was expected to decrease, while diffuse multipath was expected to increase. As a result, the expected effect of RMS surface roughness on track angle error was unknown.

Facet Size. Initial consideration was given to varying the facet size used in the model. However, since the model was developed using 500-meter facets, there was some reservation about the validity of the model for other facet sizes (Broderson, 1979). Additionally, very small facets would require an extreme amount of computer storage space to hold a reasonably-sized terrain grid. Subsequent test runs with very large facets resulted in questionable output data (Broderson, 1979). Due to these considerations, facet size was held constant at 1650 feet square (approximately 500 meters square). Facet size remains an important parameter, and once the model has been verified for other sizes, facet size should be included in the validation testing.

Small-Angle Approximations. The effect of using small-angle approximations in the model was considered for evaluation, but rejected. The effect of using them was considered negligible compared to the effect of some of the other variables included in the experimental design. (Broderson, 1979). It was therefore left out in the interest of minimizing the number of parameters in the design.

Effect of Using Integer-Valued Milliradians in ITARA and IFACET. The terrain tilts were stored in the model in integer-valued milliradians, X and Y tilt are stored in an array called ITARA. Later in the program they are converted to longitudinal and transverse tilt and stored in an array called IFACET. When the conversion is made, the resulting longitudinal and transverse tilts are in real form, but are truncated to integers in the assignment to IFACET. The maximum error introduced by truncating to integer milliradians is just less than 1 MR, or less than approximately 1/18 degree. Considering that the terrain slope was determined from a plane that was least-squares fit to the terrain, the variance in the slope of that fitted plane will in all likelihood be greater than one milliradian. Thus, concern about a one milliradian truncation error later in the program is unwarranted, and so its effect was not evaluated. An example is discussed in the next section, however, where the truncation error did cause some erroneous responses.

From this discussion, then, the only variables excluded from evaluation that were expected to significantly affect the model are the angle off boresight, the facet size, and sea versus land terrain.

Model Implementation

This section describes the procedures followed in implementing the Georgia Tech Model independent of TAC ZINGER. Some discussion of problems encountered in the process is included for the benefit of future users of the model.

The version of the model used was taken directly from a version of TAC ZINGER that was being used in July 1979. It corresponded almost exactly to the version of the model provided in the documentation. The differences between the two were due primarily to differences in the computer systems used. To implement the model independently, a driver program was built to call the model's subroutines and to supply information normally provided by TAC ZINGER. A few modifications were made to the model itself to facilitate the use of the model with the experimental design. Care was taken to insure that those changes did not cause the model to improperly perform its calculations. In most cases, cards from the original program that were changed or deleted were "commented out," instead of being physically removed from the deck. This provided a quick and easy comparison of the thesis program with the original. Comment cards were liberally added to the subroutines to clarify the program flow.

The driver program was designed to allow changes in parameters to be made in a simple, direct way. The driver is self-documenting (Appendix B), but its basics will be briefly discussed here. First of all, to simplify input, the NAMELIST capability of CDC FORTRAN IV EXTENDED was used. Using NAMELIST for input allowed the parameter values for each run to be input in any order, with each value directly identified by name. Parameters that did not need to be changed often were set in the program itself. The program as shown is in the form used with the experimental design. Provision is also made for a straight and level flight path to be used. The driver specifies the output from the program as an "echo" listing of the NAMELIST values, plus the elevation and azimuth angle errors and the modified radar cross-section. These last three are also written to a storage file for later analysis.

The changes in MULTIN were made to allow the terrain and radar characteristics to be passed from the driver instead of being assigned in MULTIN. The possibility of loading terrain data in ITARA in reverse order (X,Y instead of Y,X) has been discussed previously. To highlight the trouble spot in the code, comments were added and the order of the I and J subscripts were reversed from the original code.

MULTIP includes several changes from the original code. First, when the model is used in TAC ZINGER, the multipath reflection coefficients and the clutter patch cross-section are not calculated on every call to MULTIP. To conserve execution time, the calculations are performed only every"IRAN" calls to MULTIP, where IRAN is a constant dependent on the facet size. IRAN is calculated in MULTIN and passed via a common block to MULTIP. For the default facet size of 1650 feet square, the multipath and clutter calculations are performed once every twelfth call to MULTIP. This procedure is justified when the model is called frequently over the target's flight path, assuming the terrain does not change greatly between calls. For this research, though, MULTIP was not called as the target progressed along a flight path, and so the calculations needed to be performed on every call. To allow this to be done required the deletion of only one line of code. The change is documented in the program listing in Appendix B. The deletion of that line has no effect on the programs other than to allow the multipath and clutter calculations to be made on every call to MULTIP.

A second change to MULTIP was provided by Georgia Tech. The model, as designed, did not produce realistic multipath results for the default terrain (a flat plane with RMS surface roughness of 10 cm) (Stuk, 1979b). To remedy the situation, Georgia Tech added several lines of code to MULTIP, which were subsequently added to the model used here. One line of code in the original version of the change sent from Georgia Tech was in error, and resulted in the specular multipath calculations never being made. The error was detected when the change was implemented for this research. Georgia Tech was notified, and they in turn notified AF/SA so the correction could be disseminated to TAC ZINGER users.

A third change to the original program was also due to a programming error in the code. As in the previous case, this error was also detected during the debugging/verification of the model for this research. The erroneous code was in MULTIP, just prior to the end of the clutter section and read

```
IF (ABS(FAC1-1.).LT.001)GOTO 989.
```

The apparent intent was to compare ABS(FAC1-1.) to .001, but in the form shown it was actually compared to the integer 1. When Georgia Tech was notified, they confirmed that the original code was in error and notified AF/SA. The appropriate change was made here prior to testing the model.

The final major change to the program was made so that the effect of sloping terrain could be properly evaluated. To determine that effect, the facets used in calculating multipath and clutter needed to be tilted. The problem arose when just those facets could not be tilted. Looking at the terrain from a cross-sectional view along the longitudinal axis will illustrate the problem and the reason for the change (Figure 10).

The terrain influences the system's response by supporting multipath or clutter or both. For clutter to be supported, the facets identified for use in the clutter evaluation (Figure 11a) must be visible from the radar antenna. If those facets are hidden, or shadowed, by intervening terrain, there can be no clutter return (Figure 11b). For multipath, any facet along the longitudinal axis may support multipath provided it has the proper slope and is not shadowed from the antenna or from the

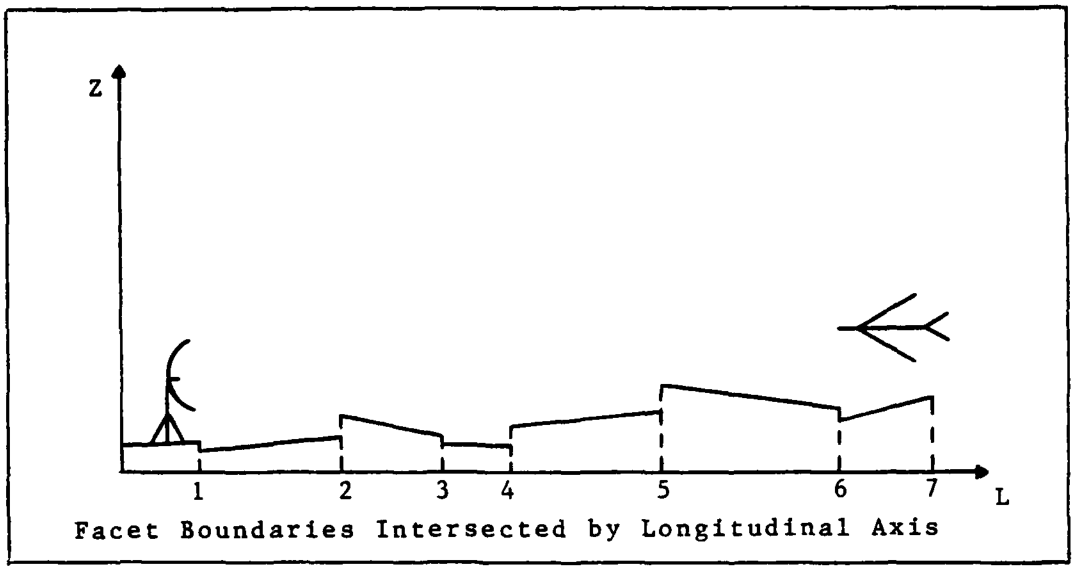

Figure 10. Terrain Cross-section Along Longitudinal Axis

aircraft (Figures 12a and 12b). Some shadowing is usually present in most applications due to the varying slope of the actual terrain. For the present application, however, the slope cannot be allowed to vary as it would with actual terrain. To evaluate its effect, the experimental design requires the slope on each run to be held constant. A real-life representation of this requirement would have only those facets tilted that could support multipath or clutter. However, since those facets are unknown prior to running the program, every facet in the grid must have the same tilt. The resulting terrain is shown in Figure 13.

The problem with this terrain representation is that only the first facet is visible from the antenna; the rest are shadowed by it. From this geometry, the model would properly determine that only the first facet could support multipath, and then only when the target was in the proper

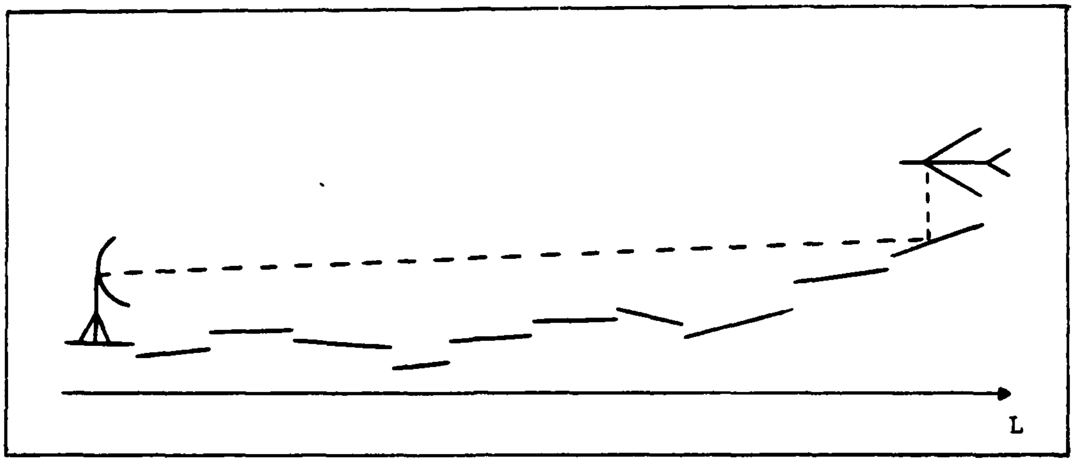

Figure 11a. Clutter Facets Visible from Site

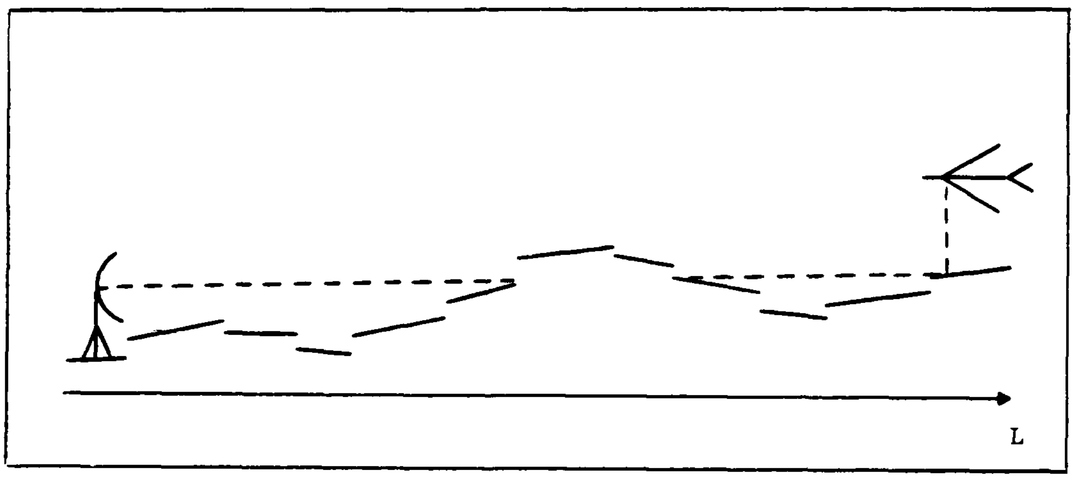

Figure 11b. Clutter Facets Shadowed by Terrain

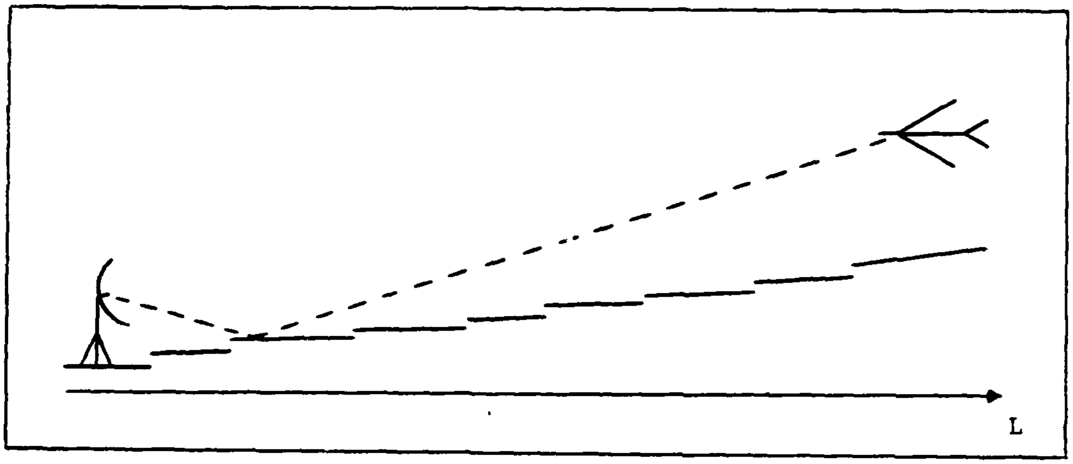

Figure 12a. All Candidate Facets for Multipath Visible from Both Site and Target

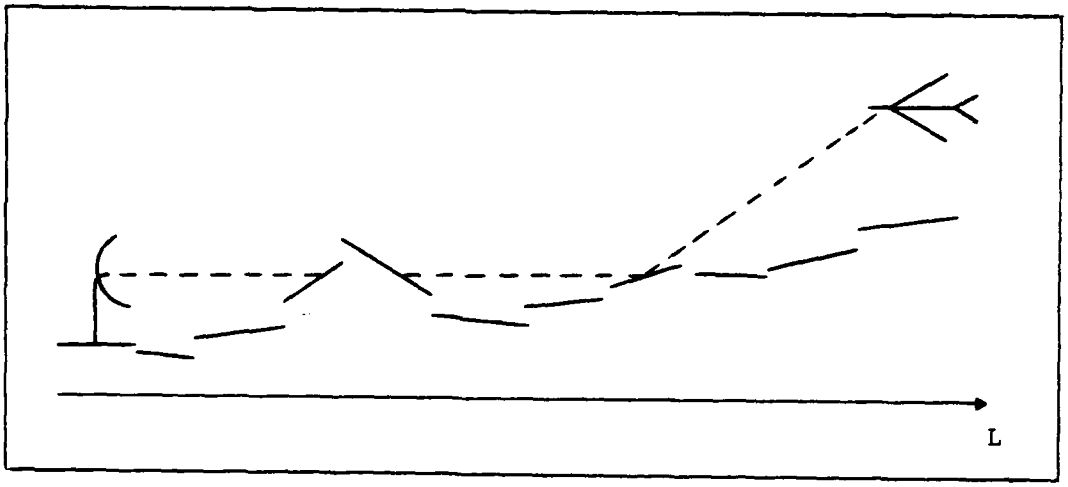

Figure 12b. Shadowing of Candidate Facets for Multipath

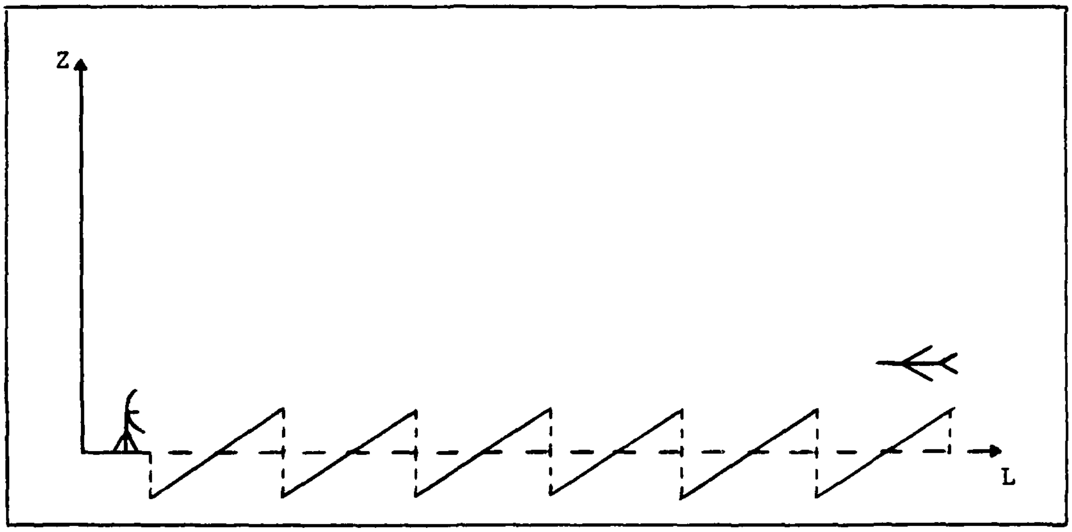

Figure 13. Terrain Facets with Equal Slope and Constant Center Heights

position. Shadowing of the clutter cell is tested separately, and will be discussed shortly.

The solution to the problem was to simply delete the check for multipath shadowing from the program, thus making intervening facets "invisible." As a result, only the proper facets are used, one at a time, for the multipath calculations. This solution is equivalent to changing the slopes and elevations of the intervening facets to values that would prevent them from shadowing the desired facets or contributing to multipath themselves. To effect this change in the program, a parameter called NOSHAD was added to the parameters in the FACET call. Its use is documented in the program.

As mentioned, shadowing of the clutter cell is tested separately from shadowing for multipath. First of all, the procedure just discussed for multipath is used in the original program for clutter. That is, the facets between the site and the clutter cell are tested as though they are invisible. Thus, based on the geometry in Figure 13 alone, the program would not determine that the clutter patch was shadowed. The test is based on another parameter, ZMASK, that is passed from TAC ZINGER in the call to MULTIP. AMASK is the maximum height of terrain along the ground line between the site and the clutter patch (Stuk, 1979b). The test simply compares the height of the terrain directly beneath the target to AMASK. If AMASK is greater, no clutter is calculated.

Unless some adjustment is made for site elevation when ZMASK is calculated in TAC ZINGER, this test for shadowing is not accurate. In the runs for this research, AMASK was set to zero, so that the clutter patch would always be visible from the site.

Finally, two other minor changes were made to the program. First, the calls to subroutine CVTXY and the corresponding conversions from meters to TSCALE units were deleted from FACET. The reason for using them is documented in the program listing in Appendix B. Although not a part of the Georgia Tech Model, CVTXY is included in Appendix B for reference. Second, the CDC uniform random number generator function RANF was used in subroutine RICE in place of UNIRAN in the original program.

The many problems encountered in implementing the model had not been expected when the research began, and so no systematic method of verification was developed or followed. Throughout the research, though, reasonable parameter values and expected results were obtained from Georgia Tech (Stuk, 1969b; Tuley, 1979b) and Major Broderson (1979). When the results were not in line with the expectations, the computer code was traced through to determine what error had been made. This trial-and-error process was time-consuming and inefficient, and most of the problems encountered could have been avoided if the documentation had been more complete. As could best be determined from such a process, the model was running properly when the data from the experimental design was gathered for evaluation.

Experimental Design Used

As noted in the previous chapter, a 2_{IV}^{16-11} fractional factorial design was chosen to be used for this experimentation. It was chosen after initial investigation of the model parameters indicated that between 12 and 18 parameters would need to be varied. Subsequent research narrowed the number to 16, as discussed earlier in this chapter.

The design used was discovered in Box and Hunter's (1961a:339-341) paper on fractional factorial designs. In that paper, the design was specified as a set of eleven generators (Table IV). The specific combinations of high/low parameter values that resulted from carrying out the

TABLE IX

GENERATORS OF THE 2^{16-11} FRACTIONAL FACTORIAL DESIGN
(Box and Hunter, 1961a:340)

$\vec{1}$	$\vec{2}$	$\vec{3}$			$\vec{6}$
$\vec{1}$	$\vec{2}$		$\vec{4}$		$\vec{7}$
$\vec{1}$	$\vec{2}$			$\vec{5}$	$\vec{8}$
$\vec{1}$		$\vec{3}$	$\vec{4}$		$\vec{9}$
$\vec{1}$		$\vec{3}$		$\vec{5}$	$\overrightarrow{10}$
$\vec{1}$			$\vec{4}$	$\vec{5}$	$\overrightarrow{11}$
	$\vec{2}$	$\vec{3}$	$\vec{4}$		$\overrightarrow{12}$
	$\vec{2}$	$\vec{3}$		$\vec{5}$	$\overrightarrow{13}$
	$\vec{2}$		$\vec{4}$	$\vec{5}$	$\overrightarrow{14}$
		$\vec{3}$	$\vec{4}$	$\vec{5}$	$\overrightarrow{15}$
$\vec{1}$	$\vec{2}$	$\vec{3}$	$\vec{4}$	$\vec{5}$	$\overrightarrow{16}$

operations implied by Table IX are shown in Table X. Since 2^{k-p}_{IV} factorials require 2 k runs, 32 runs were needed.

To develop the design from the generators given, the levels of the first five factors are listed in standard order. The remaining columns ($\vec{6}$ through $\overrightarrow{16}$) are generated by multiplying together the columns designated by the appropriate generators. For example,

$$\vec{6} = \vec{1}\vec{2}\vec{3} = \begin{array}{cccc} (-)(-)(-) & = & (-) \\ (+)(-)(-) & = & (+) \\ (-)(+)(-) & = & (+) \\ \vdots\ \ \vdots\ \ \vdots & & \vdots \\ (+)(+)(+) & = & (+) \end{array} \tag{14}$$

TABLE X

PRINCIPAL FRACTION 2^{16-11}_{IV} DESIGN

	Factor															
						$\vec{6}$	$\vec{7}$	$\vec{8}$	$\vec{9}$	$\vec{10}$	$\vec{11}$	$\vec{12}$	$\vec{13}$	$\vec{14}$	$\vec{15}$	$\vec{16}$
Run	$\vec{1}$	$\vec{2}$	$\vec{3}$	$\vec{4}$	$\vec{5}$	$\vec{1}\vec{2}\vec{3}$	$\vec{1}\vec{2}\vec{4}$	$\vec{1}\vec{2}\vec{5}$	$\vec{1}\vec{3}\vec{4}$	$\vec{1}\vec{3}\vec{5}$	$\vec{1}\vec{4}\vec{5}$	$\vec{2}\vec{3}\vec{4}$	$\vec{2}\vec{3}\vec{5}$	$\vec{2}\vec{4}\vec{5}$	$\vec{3}\vec{4}\vec{5}$	$\vec{1}\vec{2}\vec{3}\vec{4}\vec{5}$
1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	+	-	-	-	-	+	+	+	+	+	+	-	-	-	-	+
3	-	+	-	-	-	+	+	+	-	-	-	+	+	+	-	+
4	+	+	-	-	-	-	-	-	+	+	+	+	+	+	-	-
5	-	-	+	-	-	+	-	-	+	+	-	+	+	-	+	+
6	+	-	+	-	-	-	+	+	-	-	+	+	+	-	+	-
7	-	+	+	-	-	-	+	+	+	+	-	-	-	+	+	-
8	+	+	+	-	-	+	-	-	-	-	+	-	-	+	+	+
9	-	-	-	+	-	-	+	-	+	-	+	+	-	+	+	+
10	+	-	-	+	-	+	-	+	-	+	-	+	-	+	+	-
11	-	+	-	+	-	+	-	+	+	-	+	-	+	-	+	-
12	+	+	-	+	-	-	+	-	-	+	-	-	+	-	+	+
13	-	-	+	+	-	+	+	-	-	+	+	-	+	+	-	-
14	+	-	+	+	-	-	-	+	+	-	-	-	+	+	-	+
15	-	+	+	+	-	-	-	+	-	+	+	+	-	-	-	+
16	+	+	+	+	-	+	+	-	+	-	-	+	-	-	-	-
17	-	-	-	-	+	-	-	+	-	+	+	-	+	+	+	+
18	+	-	-	-	+	+	+	-	+	-	-	-	+	+	+	-
19	-	+	-	-	+	+	+	-	-	+	+	+	-	-	+	-
20	+	+	-	-	+	-	-	+	+	-	-	+	-	-	+	+
21	-	-	+	-	+	+	-	+	+	-	+	+	-	+	-	-
22	+	-	+	-	+	-	+	-	-	+	-	+	-	+	-	+
23	-	+	+	-	+	-	+	-	+	-	+	-	+	-	-	+
24	+	+	+	-	+	+	-	+	-	+	-	-	+	-	-	-
25	-	-	-	+	+	-	+	+	+	+	-	+	+	-	-	-
26	+	-	-	+	+	+	-	-	-	-	+	+	+	-	-	+
27	-	+	-	+	+	+	-	-	+	+	-	-	-	+	-	+
28	+	+	-	+	+	-	+	+	-	-	+	-	-	+	-	-
29	-	-	+	+	+	+	+	+	-	-	-	-	-	-	+	+
30	+	-	+	+	+	-	-	-	+	+	+	-	-	-	+	-
31	-	+	+	+	+	-	-	-	-	-	-	+	+	+	+	-
32	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Using this design and assuming three-factor and higher order interactions are negligible, independent estimates of the effects of factors 1 through 16 can be obtained. Sixteen additional effects can also be obtained: the grand mean, or average effect of all of the factors; and the 15 combinations of two-factor interactions shown in Table XI (Box and Hunter, 1961a:340). Each row in Table XI shows the two-factor interactions that are confounded with each other. For example, the effect of two-factor interaction $\vec{1}\vec{2}$ cannot be separated from the effects of the two-factor interactions $\overrightarrow{15}\ \overrightarrow{16}$, $\vec{3}\ \vec{6}$, $\vec{4}\ \vec{7}$, $\vec{5}\ \vec{8}$, $\vec{9}\ \overrightarrow{12}$, $\overrightarrow{10}\ \overrightarrow{13}$, and $\overrightarrow{11}\ \overrightarrow{14}$.

Since little was known about which factors would interact strongly with others, the 16 parameters were arbitrarily assigned to factors 1 through 16 as shown in Table XII.

In the initial run, several factors were found to have significant interactions with each other when evaluated against elevation angle error. Those parameters were range, target elevation, site elevation, and radar pulsewidth. As can be seen in Table XI, several of the two-factor interactions between these parameters are confounded with each other. For example, the interaction between range and target elevation ($\vec{1}\ \vec{2}$) is confounded with the interaction between antenna elevation and radar pulsewidth ($\vec{4}\ \vec{7}$). To get a better estimate of the interactions between the four variables, the parameters were rearranged so the two-factor interactions between those four variables

TABLE XI

COMBINATIONS OF TWO-FACTOR INTERACTIONS IN THE 2_{IV}^{16-11} DESIGN

(Box and Hunter, 1961a:340)

$\vec{1}\,\vec{2}$	+	$\vec{15}\,\vec{16}$	+	$\vec{3}\,\vec{6}$	+	$\vec{4}\,\vec{7}$	+	$\vec{5}\,\vec{8}$	+	$\vec{9}\,\vec{12}$	+	$\vec{10}\,\vec{13}$	+	$\vec{11}\,\vec{12}$
$\vec{1}\,\vec{3}$	+	$\vec{2}\,\vec{6}$	+	$\vec{14}\,\vec{16}$	+	$\vec{4}\,\vec{9}$	+	$\vec{5}\,\vec{10}$	+	$\vec{11}\,\vec{15}$	+	$\vec{7}\,\vec{12}$	+	$\vec{8}\,\vec{13}$
$\vec{1}\,\vec{4}$	+	$\vec{2}\,\vec{7}$	+	$\vec{3}\,\vec{9}$	+	$\vec{13}\,\vec{16}$	+	$\vec{5}\,\vec{11}$	+	$\vec{6}\,\vec{12}$	+	$\vec{10}\,\vec{15}$	+	$\vec{8}\,\vec{14}$
$\vec{1}\,\vec{5}$	+	$\vec{2}\,\vec{8}$	+	$\vec{3}\,\vec{10}$	+	$\vec{4}\,\vec{11}$	+	$\vec{12}\,\vec{16}$	+	$\vec{6}\,\vec{13}$	+	$\vec{7}\,\vec{14}$	+	$\vec{9}\,\vec{15}$
$\vec{1}\,\vec{6}$	+	$\vec{2}\,\vec{3}$	+	$\vec{14}\,\vec{15}$	+	$\vec{4}\,\vec{12}$	+	$\vec{5}\,\vec{13}$	+	$\vec{11}\,\vec{16}$	+	$\vec{7}\,\vec{9}$	+	$\vec{8}\,\vec{10}$
$\vec{1}\,\vec{7}$	+	$\vec{2}\,\vec{4}$	+	$\vec{3}\,\vec{12}$	+	$\vec{13}\,\vec{15}$	+	$\vec{5}\,\vec{14}$	+	$\vec{6}\,\vec{9}$	+	$\vec{10}\,\vec{16}$	+	$\vec{8}\,\vec{11}$
$\vec{1}\,\vec{8}$	+	$\vec{2}\,\vec{5}$	+	$\vec{3}\,\vec{13}$	+	$\vec{4}\,\vec{14}$	+	$\vec{12}\,\vec{15}$	+	$\vec{6}\,\vec{10}$	+	$\vec{7}\,\vec{11}$	+	$\vec{9}\,\vec{16}$
$\vec{1}\,\vec{9}$	+	$\vec{2}\,\vec{12}$	+	$\vec{3}\,\vec{4}$	+	$\vec{10}\,\vec{11}$	+	$\vec{5}\,\vec{15}$	+	$\vec{6}\,\vec{7}$	+	$\vec{13}\,\vec{14}$	+	$\vec{8}\,\vec{16}$
$\vec{1}\,\vec{10}$	+	$\vec{2}\,\vec{13}$	+	$\vec{3}\,\vec{5}$	+	$\vec{4}\,\vec{15}$	+	$\vec{12}\,\vec{14}$	+	$\vec{6}\,\vec{8}$	+	$\vec{7}\,\vec{16}$	+	$\vec{9}\,\vec{11}$
$\vec{1}\,\vec{11}$	+	$\vec{2}\,\vec{14}$	+	$\vec{3}\,\vec{15}$	+	$\vec{4}\,\vec{5}$	+	$\vec{9}\,\vec{10}$	+	$\vec{6}\,\vec{16}$	+	$\vec{7}\,\vec{8}$	+	$\vec{12}\,\vec{13}$
$\vec{1}\,\vec{12}$	+	$\vec{2}\,\vec{9}$	+	$\vec{3}\,\vec{7}$	+	$\vec{4}\,\vec{6}$	+	$\vec{5}\,\vec{16}$	+	$\vec{10}\,\vec{14}$	+	$\vec{8}\,\vec{15}$	+	$\vec{11}\,\vec{13}$
$\vec{1}\,\vec{13}$	+	$\vec{2}\,\vec{10}$	+	$\vec{3}\,\vec{8}$	+	$\vec{4}\,\vec{16}$	+	$\vec{5}\,\vec{6}$	+	$\vec{11}\,\vec{12}$	+	$\vec{7}\,\vec{15}$	+	$\vec{9}\,\vec{14}$
$\vec{1}\,\vec{14}$	+	$\vec{2}\,\vec{11}$	+	$\vec{3}\,\vec{16}$	+	$\vec{4}\,\vec{8}$	+	$\vec{5}\,\vec{7}$	+	$\vec{6}\,\vec{15}$	+	$\vec{9}\,\vec{13}$	+	$\vec{10}\,\vec{12}$
$\vec{1}\,\vec{15}$	+	$\vec{2}\,\vec{16}$	+	$\vec{3}\,\vec{11}$	+	$\vec{4}\,\vec{10}$	+	$\vec{5}\,\vec{9}$	+	$\vec{6}\,\vec{14}$	+	$\vec{7}\,\vec{13}$	+	$\vec{8}\,\vec{12}$
$\vec{1}\,\vec{16}$	+	$\vec{2}\,\vec{15}$	+	$\vec{3}\,\vec{14}$	+	$\vec{4}\,\vec{13}$	+	$\vec{5}\,\vec{12}$	+	$\vec{6}\,\vec{11}$	+	$\vec{7}\,\vec{10}$	+	$\vec{8}\,\vec{9}$

TABLE XII

PARAMETER ORDER FOR INITIAL RUN

Factor	Parameter	Factor	Parameter
1	Range	9	Clutter Parameter A
2	Target Elevation	10	Clutter Parameter B
3	Radar Cross-section of Target	11	Clutter Parameter C
4	Antenna Elevation	12	Clutter Parameter D
5	Radar Frequency	13	Dielectric Constant
6	Radar Beamwidth	14	RMS Surface Roughness
7	Radar Pulsewidth	15	Longitudinal Tilt
8	Sub-clutter Visibility	16	Transverse Tilt

were not confounded with each other. This was done by simply switching the parameters associated with $\vec{7}$ and $\vec{13}$, so that $\vec{7}$ became the dielectric constant and $\vec{13}$ became the radar pulsewidth.

Each of the 32 runs was replicated 30 times to get a reasonably accurate estimate of the mean and variance of the response for each run. Each run used the same random number stream to reduce variance between runs, but each replication within a run used a different set of random numbers.

Data Analysis

The data was analyzed using a stepwise multiple regression technique available in the Statistical Package for the Social Sciences (SPSS). Regressions were run for all three response variables (elevation and azimuth track angle errors, and modified radar cross-section). Thirty-one terms were included as independent variables: the 16 factors in the experimental design, plus one term each for the 15 combinations of two factor interactions shown in Table XI.

The stepwise regression technique introduced the independent variables into the right-hand side of the regression equation one at a time, and, if necessary, also deleted variables one at a time. The addition or deletion of variables was based on the significance level at which the null hypothesis

$$H_0: \quad \beta_i = 0, \tag{15}$$

where β_i is the coefficient of the variable, is rejected in favor of the alternate hypothesis

$$H_A: \quad \beta_i \neq 0. \tag{16}$$

The significance was tested using the F statistic. The most significant variable, i.e., the one with the largest F statistic, was added to the equation each iteration, provided that F statistic was greater than a predetermined value. If, once a variable was in the equation, its significance dropped below another preset value, that variable was deleted from the equation. The final equation thus contained only variables whose coefficients were different from zero at a predetermined statistical significance level.

The results of the first analysis were used to rearrange the order of the variables in the experimental design, so that better estimates of the two-factor interactions could be obtained. The next chapter discusses the results of these analyses.

V. Data Analysis

The results of the data analysis procedures described in Chapter IV are presented in Tables XIVa, XIVb, XVa, and XVb. The tables present results for elevation track angle error and modified radar cross-section. Data on azimuth track angle error was not included for reasons given later. As mentioned in Chapter IV, the principal emphasis in the analysis was on elevation track angle error.

The factors in the design matrix and their corresponding parameters for the initial run are shown in Table XIIIa. Tables XIVa and XIVb are the results of the analysis of the data generated by that design. The interaction terms shown are simply the first terms from each set of confounded two-factor interactions shown in Table XI. The variables are listed in the order they entered the equation. $\hat{\beta}$ is the least-squares estimator of the coefficient of each effect. The change in R^2 is the change due to adding the variable to those already in the equation.

Table XIVa shows the 17 effects whose coefficients were statistically different from zero, based on a minimum F-statistic value of 4.0. Of those 17, 7 were significantly more important than the others: range, target and antenna elevations, radar pulsewidth, and the interactions between range and each of the other three. Their relative

TABLE XIIIa

PARAMETER VALUES AND ORDER IN ORIGINAL DESIGN MATRIX

Factor	Parameter	High Value	Low Value
1	Range	50000 FT	3000 FT
2	Target Elevation	1000 FT	100 FT
3	Radar Cross-Section of Target	.01 $METER^2$	10 $METERS^2$
4	Antenna Elevation	100 FT	6 FT
5	Radar Frequency	9 GHz	16 GHz
6	Radar Beamwidth	.8 DEGREES	1.2 DEGREES
7	Radar Pulsewidth	100 NANOSEC	300 NANOSEC
8	Sub-Clutter Visibility	1 dB	50 dB
9	Clutter Parameter A	.079	.0045
10	Clutter Parameter B	1.5	.83
11	Clutter Parameter C	.012	.0013
12	Clutter Parameter D	2.3	0.0
13	Dielectric Constant	20.0 + 2.41	2.0
14	RMS Surface Roughness	20 cm	500 cm
15	Longitudinal Tilt	-20 DEG	0 DEG
16	Transverse Tilt	-20 DEG	0 DEG

TABLE XIIIb

INTERACTION DEFINITIONS FOR ANALYSIS

Interaction	Effects	Interation	Effects
1X2	$\vec{1}$ $\vec{2}$	1X9	$\vec{1}$ $\vec{9}$
1X3	$\vec{1}$ $\vec{3}$	1X10	$\vec{1}$ $\vec{10}$
1X4	$\vec{1}$ $\vec{4}$	1X11	$\vec{1}$ $\vec{11}$
1X5	$\vec{1}$ $\vec{5}$	1X12	$\vec{1}$ $\vec{12}$
1X6	$\vec{1}$ $\vec{6}$	1X13	$\vec{1}$ $\vec{13}$
1X7	$\vec{1}$ $\vec{7}$	1X14	$\vec{1}$ $\vec{14}$
1X8	$\vec{1}$ $\vec{8}$	1X15	$\vec{1}$ $\vec{15}$
		1X16	$\vec{1}$ $\vec{16}$

TABLE XIVa

ANALYSIS OF ELEVATION ERRORS WITH PULSEWIDTH AND DIELECTRIC CONSTANT IN ORIGINAL ORDER

MULTIPLE R = .99333	OVERALL F = 4110.7
R^2 = .98670	DEGREES OF FREEDOM
ADJUSTED R^2 = .9864	REGRESSION = 17
	RESIDUAL = 942
	SIGNIFICANCE < .001

VARIABLES IN THE EQUATION

Variable	$\hat{\beta}$	F	Significance	Change in R^2
Target Elevation	-.00632	41580	<.001	.58709
Antenna Elevation	.00248	6412	<.001	.09054
1X7	-.00245	6280	<.001	.08867
1X2	-.00195	3980	<.001	.05621
Range	.00192	3856	<.001	.05446
Pulsewidth	.00191	3791	<.001	.05353
1X4	-.00188	3690	<.001	.05210
1X3	.00019	37.8	<.001	.00053
Clutter Parameter C	-.00019	37.8	<.001	.00053
1X11	-.00019	37.8	<.001	.00053
Radar Cross-Section	.00019	37.8	<.001	.00053
1X14	.00018	32.7	<.001	.00046
Beamwidth	-.00018	32.7	<.001	.00046
1X6	-.00018	32.7	<.001	.00046
Surface Roughness	-.00018	32.7	<.001	.00046
1X10	-.00006	4.34	<.038	.00006
Clutter Parameter	-.00006	4.33	<.038	.00006

VARIABLES NOT IN THE EQUATION

Variable	Partial F	Significance
Dielectric Constant	1.57	.210
1X13	1.57	2.11
Clutter Parameter D	.747	.388
1X8	.747	.388
Sub-clutter Visibility	.746	.388
1X12	.746	.388
Longitudinal Tilt	.522	.470
1X15	.521	.471
Frequency	.297	.586
1X9	.297	.586
1X5	.296	.587
Clutter Parameter A	.296	.587
Transverse Tilt	.072	.788
1X16	.072	.789

TABLE XIVb

ANALYSIS OF MODIFIED RADAR CROSS-SECTION WITH PULSEWIDTH AND DIELECTRIC CONSTANT IN ORIGINAL ORDER

MULTIPLE R^2	= .98952	OVERALL F	= 3707
R^2	= .97915	DEGREES OF FREEDOM	
Adjusted R	= .97889	REGRESSION	= 12
		RESIDUAL	= 947
		SIGNIFICANCE	< .001

VARIABLES IN THE EQUATION

Variable	$\hat{\beta}$	F	Significance	Change in R^2
Radar Cross-section	14.45	44047	<.001	.98471
1X3	-.549	63.5	<.001	.00140
Clutter Parameter C	.548	63.4	<.001	.00140
1X11	.548	63.4	<.001	.00140
1X6	.514	55.7	<.001	.00123
1X14	-.514	55.7	<.001	.00123
Surface Roughness	-.514	55.6	<.001	.00122
Beamwidth	.513	55.6	<.001	.00122
1X8	-.149	4.68	.031	.00010
Clutter Parameter D	.149	4.68	.031	.00010
1X12	.149	4.67	.031	.00010
Sub-clutter Visibility	.149	4.67	.031	.00010

VARIABLES NOT IN THE EQUATION

Variable	Partial F	Significance
Antenna Elevation	3.699	.055
Clutter Parameter B	3.699	.055
1X4	3.699	.055
1X10	3.699	.055
Dielectric Constant	3.295	.070
1X7	3.295	.070
Pulsewidth	3.290	.070
1X13	3.290	.070
Frequency	2.961	.086
1X9	2.961	.086
Clutter Parameter A	2.956	.086
1X5	2.956	.086
Target Elevation	1.729	.189
Transverse Tilt	1.721	.190
1X16	1.717	.190
1X2	1.709	.191
Range	1.323	.250
1X15	1.312	.252
Longitudinal Tilt	1.309	.253

importance was evaluated on the basis of three things: change in R^2, the values of their F-statistics, and the values of their coefficients.

The overall R^2, or fraction of the variance in elevation angle error explained by the regression equation, is an extremely high .98670. However, the seven effects noted above alone account for .98259 of the variance. As Table XIVa shows, those seven each explain at least an additional 5 percent of the variance, while the remainder combined explain less than one-half a percent.

The relative values of the F-statistics of the first seven compared to the remainder also point out the importance of those variables. The F-statistics of the first seven are on the order of at least 100 times greater than those of the other ten.

Recalling the discussion in Chapter III on the interpretation of analysis results, the $\hat{\beta}$'s also provide a good estimate of the relative importance of the variables. This is a result of standardizing the range of each variable as the range between its extreme high and low values. The coefficients of the seven variables under consideration are at least ten times larger than the coefficients of any of the others. This is a direct indication that the seven have at least ten times the effect on the response as any of the others.

Thus far, the analysis has supported the original hypothesis that only few variables and their interactions are really important in the model. The design used provided estimates of four important parameters and three estimates of two-factor interactions between them. However, six two-factor interactions were possible between those four factors. The remaining three were confounded with the three reported. If those confounded interactions could be separated, better estimates of all six would result.

As described in Chapter IV, the positions of the pulsewidth and dielectric constant in the design matrix were switched to separate those confounded interactions. The data produced after switching the variable positions are shown in Table XVa and XVb.

If the first run had been accurately described by the regression analysis, this second run should look very similar. The results shown are disturbingly different from the first results. First, and probably most important, pulsewidth did not have a significant effect this time. Instead, the dielectric constant and its interactions came into the equation, and pulsewidth did not. Second, only seven variables entered the equation at all, instead of 17. Those that did enter corresponded exactly with the <u>positions</u> in the design matrix of the variables that entered on the first run. That is, columns $\vec{1}$, $\vec{2}$, $\vec{4}$, and $\vec{7}$. Third, the R^2 dropped from .987 to .777.

The F-statistics for each of the variables in the equation was large enough to reject the hypothesis that the coefficient of the variable was zero at a significance level

TABLE XVa

ANALYSIS OF ELEVATION ERRORS WITH PULSEWIDTH AND DIELECTRIC CONSTANT IN REVERSED ORDER

MULTIPLE R	= .88240	OVERALL F	= 478.4
R^2	= .77862	DEGREES OF FREEDOM	
ADJUSTED R^2	= .77701	REGRESSION	= 7
		RESIDUAL	= 952
		SIGNIFICANCE	< .001

VARIABLES IN THE EQUATION

Variable	$\hat{\beta}$	F	Significance	Change in R^2
Target Elevation	-.00648	1990	<.001	.46272
Antenna Elevation	.00264	330.5	<.001	.07685
1X7	-.00262	324.0	<.001	.07534
1X2	-.00212	212.5	<.001	.04942
Range	.00209	159.2	<.001	.04799
Dielectric Constant	.00175	126.3	<.001	.03366
1X4	-.00172	140.5	<.001	.03266

VARIABLES NOT IN THE EQUATION

Variable	Partial F	Significance
Clutter Parameter B	2.4	.122
1X10	2.4	.122
Pulsewidth	1.87	.172
1X13	1.87	.172
Traverse Tilt	1.4	.238
1X16	1.39	.238
Longitudinal Tilt	.938	.333
1X15	.928	.333
Clutter Parameter A	.896	.344
1X5	.896	.344
1X9	.896	.344
Frequency	.895	.344
Sub-Clutter Visibility	.777	.378
Clutter Parameter D	.777	.378
1X8	.777	.378
1X12	.777	.378
Radar Cross-Section	.043	.836
Clutter Parameter C	.043	.836
1X3	.043	.836
1X11	.043	.836
Beamwidth	.014	.906
Surface Roughness	.014	.906
1X6	.014	.906
1X14	.014	.906

TABLE XVb

ANALYSIS OF MODIFIED RADAR CROSS-SECTION WITH PULSEWIDTH AND DIELECTRIC CONSTANT IN REVERSED ORDER

MULTIPLE R	= .98888	OVERALL F	=	3490
R^2	.97789	DEGREES OF FREEDOM		
ADJUSTED R^2	.97761	REGRESSION	=	12
		RESIDUAL	=	947
		SIGNIFICANCE	<	.001

VARIABLES IN THE EQUATION

Variable	$\hat{\beta}$	F	Significance	Change in R^2
Radar Cross-Section	14.45	41461	<.001	.96817
1X3	-.553	60.7	<.001	.00142
Clutter Parameter C	.552	60.6	<.001	.00141
1X11	.552	60.6	<.001	.00141
1X6	.518	53.3	<.001	.00124
1X14	-.418	53.2	<.001	.00124
Surface Roughness	-.518	53.2	<.001	.00124
Beamwidth	.517	53.2	<.001	.00124
1X8	-.164	5.37	.021	.00013
Clutter Parameter D	.164	5.37	.021	.00013
Sub-Clutter Visibility	-.164	5.36	.021	.00013
1X12	.164	5.36	.021	.00013

VARIABLES NOT IN THE EQUATION

Variable	Partial F	Significance
Frequency	3.561	.059
1X9	3.561	.059
Clutter Parameter A	3.556	.060
1X5	3.556	.060
Antenna Elevation	3.469	.063
1X10	3.469	.063
Clutter Parameter B	3.463	.063
1X4	3.463	.063
Pulsewidth	3.125	.077
1X7	3.125	.077
Dielectric Constant	3.120	.078
1X13	3.120	.078
Target Elevation	1.864	.172
Transverse Tilt	1.856	.173
1X16	1.852	.174
1X2	1.844	.175
Range	1.480	.224
1X15	1.470	.226
Longitudinal Tilt	1.465	.226

of less than .001. For the variables not in the equation, the largest F-statistic was 2.397, with a corresponding significance of .122.

In trying to justify the apparent inconsistency in results, the most plausible solution seems to be that third or higher order interactions were not negligible. That they were negligible was, of course, one of the primary assumptions that allowed the use of a Resolution IV design. If in fact there was a large three-factor interaction between, say, range and the site and target elevations, it would have been confounded with at least one of the 31 effects calculated in the analysis. The specific effects it was confounded with could be determined only by using the technique discussed in Chapter III. Each of the 2047 defining relations for the design would have to be multiplied by $\vec{1}\vec{2}\vec{3}$, but first the 2047 defining relations would have to be generated from the 11 generating relations listed in Table IX. Due to the work involved for the value of the results obtained, those confounded effects were not determined. However, since range and target and site elevations were important in both runs, it is possible that any significant higher-order interactions would contain one or more of those variables.

To determine if there are significant interactions between those three variables, a 2^3 full-factorial design could be run, and would require only 8 runs. A 2^5 full-factorial design could be run in only 32 runs. This would allow up to five-factor interactions to be calculated. Since the effects of pulsewidth and dielectric constant are in question, they would seem logical choices for the additional two factors. Since other two-factor interactions did not appear significant in the 2_{IV}^{16-11} design, there would probably be no need for further runs beyond the 2^5 design.

With perfect tracking (i.e., angle off boresight equals zero), only four runs (2, 14, 22, 26) in both cases produced azimuth track angle errors that were not 0.0. As a result, the regression equation for the original run explained only 22 percent of the variance. For the revised run, none of the variables had a coefficient that was statistically different from zero at an F value of 3.7.

Finally, 12 of the 32 runs in both cases reported an elevation angle error of exactly zero. Those runs were 3, 5, 7, 8, 12, 15, 17, 20, 27, 31, and 32. When the model reports angle errors of exactly zero, it is usually the result of situations for which the errors induced by multipath and/or clutter terms are expected to be very small. From the equations in the code, the primary factor in such situations is the angle off boresight. In the runs made with the revised parameter order, no other factors appear to be common to the situations that produced zero angle errors.

Since there was no variation in the elevation angle errors among the 12 runs that reported values of 0.0, the data analysis was not able to distinguish between the effect of certain parameters that were varied during those runs. This is evidenced in Table XIVa by $\hat{\beta}$'s, F statistic values and changes in R^2 that are the same for more than one parameter. For instance, the analysis could not distinguish between clutter parameter C, 1 X 11, radar cross-section and 1 X 14. All four have the same $\hat{\beta}$'s, F-statistics and changes in R^2. As the data stands, those variables that have the same statistical values must be interpreted as having the same effect on the response. This is a highly unlikely situation, and suggests that there are factors relevant to the system that were not considered. The factor that should be given primary consideration for inclusion is angle off boresight. With it included, much better results would be expected.

VI. Summary and Conclusions

The goal of this thesis was to develop insight into appropriate methodology to validate computer simulations used in Operational Test and Evaluation. The conclusions reached in pursuing this goal fall into three categories: those resulting from the background investigations into validation and experimental design in validation; those pertaining to the specific research conducted; and those regarding the applicability of the previous conclusions in the OT&E context.

Background Investigation

The term validation implies more than simple agreement between actual and simulated data. Perfect agreement under all circumstances can never be reached, and is not even always desired. What is desired is confidence on the part of the user that an inference he makes from simulated data is a valid inference for the actual system. Validation is the process of building that confidence up to an acceptable degree. What is acceptable depends on the purpose for using the simulation. The expected cost of accepting a lower level of confidence must be weighed against the expected cost of obtaining a higher level.

Building a high degree of confidence in the validity of a model requires the model be tested over as broad a range as possible of its expected application. Experimental design techniques accomplish this more efficiently. Even so, too many runs would be required to vary every factor over its entire range. Those factors that have the most effect on the response should be able to be identified using screening designs. A more complete design can then be used to specify the levels of those factors for use in gathering both simulated and actual data.

Such experimental designs are limited in that they provide information only from discrete points throughout the range of possible responses. For applications where the factor levels are continually changing, e.g., a missile flight path, these designs neglect the information that is available along the flight paths between the discrete points. Tytula (1978) proposes an approach that uses the information available from all along a flight path. That approach could be applied to flight paths that include the design points determined by the experimental design.

For applications which will allow the use of discontinuous changes in parameter settings, such as this research, experimental design techniques may be used directly.

Research with the Georgia Tech Model

The five objectives set down at the beginning of Chapter IV were accomplished:

1. The Georgia Tech Model was properly implemented independently of TAC ZINGER, although many more problems were encountered than expected.

2. The appropriate parameters to test for importance were identified from the model prior to testing, except the angle off-boresight. It was considered but rejected for testing as a result of a misinterpretation of the definitions of the parameters used to calculate it.

3. An appropriate experimental design, a 2^{16-11}_{IV} fractional factorial, was determined for use in screening for the important factors in the model.

4. Data was generated using the model and that design.

5. The data was analyzed using regression analysis.

The hypothesis that there were relatively few very important parameters in the model was supported. The experimental design used identified range, antenna and target elevations, and their two-factor interactions as important factors. However, the results of the data analysis indicated that there was probably at least one unidentified higher-order interaction, and that the factors included in the analysis did not adequately explain the behavior of the system. The exclusion of angle off-boresight was suspected to be the major cause of the latter problem.

The assumption that there were no significant third or higher-order interactions apparently did not hold. To further test this, data would need to be gathered with angle off-boresight included in the experimental design.

The frequency of runs of the model that result in track angle errors of exactly 0.0 when the angle off-boresight is zero is disturbing. Due to the random nature of the errors due to clutter and multipath, some disturbance from zero was expected for cases other than free space calculations (i.e., no clutter or multipath returns).

Extrapolation of Research to OT&E

The specific approach to identifying the important parameters taken in this research will probably be directly applicable to OT&E only under certain conditions. The simulation must allow data to be gathered only at particular combinations of parameters. It cannot, for instance, make efficient use of data from points along a flight path (simulated or real) other than those specified in the design.

The research resulted in two proposals that are expected to be applicable to OT&E:

1. The support given the hypothesis of there being only a few very important parameters out of many possible lends support to the feasibility of another proposal: that in collecting data for validation, varying those parameters over their expected ranges should provide adequate exploration of the system's feasible responses.

2. Tytula's (1978) approach to validation of missile systems appears to be particularly applicable to OT&E. As described, it would require very few test launches, and it explicitly incorporates decision analysis into the validation process. It appears that using that approach in conjunction with an experimental design will aid in assuring that the validation results are applicable over the full range of the system's response.

Bibliography

Barton, David K. "Low-Angle Radar Tracking," Proceedings of the IEEE, 62(6):687-704 (June 1974).

Barton, D. K. "Multipath Fluctuation Effects in Track-While-Scan Radar," 23rd Annual Tri-Service Symposium, 1977. Unclassified copy, edited. Original confidential paper.

Bauer, Michael F. and Keki B. Irani. A Simulation Model of TSS 8.1. Paper describing design and implementation of simulation model of a large-scale computing system. Ann Arbor, Michigan: Systems Engineering Laboratory, University of Michigan, February 1976. (AD A022668).

BDM/W-77-138-BR. TAC REPELLER, A Monte Carlo Simulation of Interaction Between an Array of AAA and SAM Units and Attacking Aircraft. Report presented to the Tactical Air Combat Analysis Section of the Joint National ORSA/TIMS Meeting, 1977. McLean, VA: The BDM Corporation, 7 November 1977.

Berman, Morton B. Notes on Validating/Verifying Computer Simulation Models. Paper P-4891. Santa Monica, CA: The RAND Corporation, August 1972.

Box, G. E. P. and J. S. Hunter. "The 2^{k-p} Fractional Factorial Designs, Part I," Technometrics, 3(3):311-351 (August 1961).

Box, G. E. P. and J. S. Hunter. "The 2^{k-p} Fractional Factorial Designs, Part II, Technometrics, 3(4):449-458 (November 1961).

Broderson, R. L. Analyst in Cruise Missile Testing Branch of Aerospace Systems Division (ASD/ENFX). (Personal conversations between May and September, 1979), Wright-Patterson AFB, Ohio.

Brookner, Eli. Radar Technology. Dedham, MASS: Artech House, 1977.

Cochran, William G. and Gertrude M. Cox. Experimental Designs (second edition). New York: John Wiley and Sons, Inc., 1957.

Conway, R. W. "Some Tactical Problems in Digital Simulation," Management Science, X(1):47-61 (October 1963).

Crabill, Thomas B. and Warren E. Walker. "Validation of a Police Patrol Simulation Model," Proceedings of the Eighth Annual Simulation Symposium, 229-245. Tampa FLA: Annual Simulation Symposium, 1975.

Driscoll, Thomas R. and Robert C. Stockdale. "Validation Methodology of a Digital 6-DOF Tactical Missile Simulation Model," Proceedings of the 1975 Summer Computer Simulation Conference, 1217-1222. LaJolla, CA: The Simulation Councils, Inc., 1975.

Fishman, George S. Concepts and Methods in Discrete Digital Simulation. New York: John Wiley and Sons, Inc., 1973.

Fishman, George S. and Philip J. Kiviat. "The Analysis of Simulation-Generated Time Series," Management Science, 13(7):525-557 (March 1967).

Fishman, Goerge S. and Philip J. Kiviat. Digital Computer Simulation: Statistical Considerations. Memorandum RM-5387-PR. Santa Monica, CA: The RAND Corporation, November 1967.

Forrester, Jay W. Industrial Dynamics. Cambridge, MASS: The M.I.T. Press, 1961.

Friedman, Milton. Essays in Positive Economics. Chicago: University of Chicago Press, 1953.

Garratt, Michael. "Statistical Validation of Simulation Models," Proceedings of the 1974 Summer Computer Simulation Conference, 915-927. LaJolla, CA: The Simulation Councils, Inc., 1974.

Gilmour, Peter. "A General Validation Procedure for Computer Simulation Models," The Australian Computer Journal, 5(3):127-131 (November 1973).

Golub, Jerome. "The Validation of MEDUSA," Proceedings of the 1976 Summer Computer Simulation Conference, 699-703. LaJolla, CA: The Simulation Councils, Inc., 1976.

Hermann, Charles F. "Validation Problems in Games and Simulations," Behavioral Science, 12:216-231 (May 1967).

Hicks, Charles R. Fundamental Concepts in the Design of Experiments. New York: Holt, Rinehart and Winston, Inc., 1964.

Howrey, P. and H. H. Kelejian. "Simulation versus Analytical Solutions," The Design of Computer Simulation Experiments, edited by Thomas H. Naylor. Durham, NC: Duke University Press, 1969.

Hunter, J. S. and Thomas H. Naylor. "Experimental Design," The Design of Computer Simulation Experiments, edited by Thomas H. Naylor. Durham, NC: Duke University Press, 1969.

Kerr, Donald E. (editor). Propagation of Short Radio Waves. New York: McGraw-Hill Book Company, Inc., 1951.

Kheir, Naim A. "A Validation Procedure for Missile-Systems Simulation Models," Proceedings of the 1976 Pittsburgh Conference on Modeling and Simulation, 26-27 April 1976.

Kleijnen, Jack P. C. Statistical Techniques in Simulation, II. New York: Marcel Dekker, Inc., 1975.

Kosovac, Stephen . and George E. Shortle, Jr. "Ballistic Missile Defense Simulation Validation," Proceedings of the 1976 Summer Computer Simulation Conference, 693-698. LaJolla, CA: The Simulation Councils, Inc., 1976.

Locke, Arthur S. and Collaborators. Guidance. Princeton, N.J.: D. Van Nostrand Company, Inc., 1955.

Long, Maurice W. Radar Reflectivity of Land and Sea. Lexington, MA: D. C. Heath and Company, 1975.

Lovell, Clifton C. Simulation in Field Testing. Paper P-4152. Santa Monica, CA: The RAND Corporation, August 1969.

McKenney, James L. "Critique of: 'Verification of Computer Simulation Models,'" Management Science, 14(2):B-102 to B-103 (October 1967).

McNichols, Charles W. An Introduction To: Applied Multivariate Analysis--Course Notes. Wright-Patterson AFB, Ohio: Air Force Institute of Technology, 1979.

Mendenhall, William and Richard L. Scheaffer. Mathematical Statistics with Applications. North Scituate, MASS: Duxbury Press, 1973.

Merten, Alan G. et al. Design of Computer Systems Simulation Experiments. Report for US Army Computer Systems Command. Ann ARbor, MI: The University of Michigan, 1973.

Mihram, G. Arthur. "Some Practical Aspects of Verification and Validation of Simulation Models," Operational Research Quarterly, 23(1):17-29 (March 1972).

Miller, D. R. "Model Validation Through Sensitivity Analysis," Proceedings of the 1974 Summer Computer Simulation Conference, 911-914. LaJolla, CA: The Simulation Councils, Inc., 1974.

Naylor, Thomas H. The Design of Computer Simulation Experments. Durham, N.C.: Duke University Press, 1969.

Naylor, Thomas H. and J. M. Finger. "Verification of Computer Simulation Models," Management Science, 14(2):B-92 to B-101 (October 1967).

Naylor, Thomas H., et al. Computer Simulation Techniques. New York: John Wiley and Sons, Inc., 1968.

Naylor, Thomas H., et al. "The Design of Computer Simulation Experiments," The Design of Computer Simulation Experiments, edited by Thomas H. Naylor. Durham, N.C.: Duke University Press, 1969.

Noe, J. D. and G. J. Nutt. "Validation of a Trace-Driven CDC 6400 Simulation," AFIPS Conference Proceedings, 40: 749-757 (1972).

Nolan, Richard L. "Verification/Validation of Computer Simulation Models," Proceedings of the 1972 Summer Computer Simulation Conference, 1254-1265. LaJolla, CA: The Simulation Councils, Inc., 1972.

Shannon, Robert E. Systems Simulation: The Art and Science. Englewood Cliffs, N.J.: Prentice Hall, Inc., 1975.

Schatzoff, M. and C. C. Tillman. "Design of Experiments in Simulator Validation," IBM Journal of Research and Development, 19:252-262 (May 1975).

Schlesinger, S., et al. "Developing Standard Procedures for Simulation Validation and Verification," Proceedings of the 1974 Summer Computer Simulation Conference, 927-933. LaJolla, CA: The Simulation Councils, Inc., 1974.

Schrank, William E. and Charles C. Holt. "Critique of 'Verification of Computer Simulation Models,'" Management Science, 14(2):B-104 to B-106 (October 1967).

Shannon, Robert E. Systems Simulation: The Art and the Science. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1975.

Skolnik, Merrill I. Introduction to Radar Systems. New York: McGraw-Hill, 1962.

Stuk, Stephen P. Software Documentation for TAC ZINGER Clutter and Multipath Models. Project No. A-2194. Atlanta, GA: Engineering Experiment Station, Georgia Institute of Technology, May 1979.

Stuk, Stephen P. Author of Software Documentation for TAC ZINGER Clutter and Multipath Models. (Personal telephone conversations between July and December 1979).

Theil, Henri. Applied Economic Forecasting. Chicago: Rand McNally, 1966.

Theil, Henri. Economic Forecasts and Policy. Amsterdam: North-Holland Publishing Co., 1961, 1970.

Tuley, M. T. Editor of Development and Validation of Multipath and Clutter Models for TAC ZINGER in Low Altitude Scenarios. (Briefing on above report to users of TAC ZINGER), Wright-Patterson AFB, OH, 26-27 June 1979.

Tuley, M. T. Editor of Development and Validation of Multipath and Clutter Models for TAC ZINGER in Low Altitude Scenarios. (Personal telephone conversations between August and November 1979).

Tytula, Thomas P. A Method for Validating Missile System Simulation Models. Technical Report E-78-11 for U.S. Army Missile Research and Development Command. Redstone Arsenal, AL: USAMRDC, 1978. (AD A055 689).

Van Horn, Richard. "Validation," The Design of Computer Simulation Experiments, edited by Thomas H. Naylor. Durham, N.C.: Duke University Press, 1969.

Watts, D. "Time Series Analysis," The Design of Computer Simulation Experiments, edited by Thomas H. Naylor. Dunham, N.C.: Duke University Press, 1969.

Wilk, M. B. and R. Gnanadesikan. "Probability Plotting Methods for the Analysis of Data," Biometrrika, 55(1): 1-17 (1968).

Wright, Richard D. "Validating Dynamic Models: An Evaluation of Tests of Predictive Power," Proceedings of the 1972 Summer Computer Simulation Conference, 1286-1296. LaJolla, CA: The Simulation Councils, Inc., 1972.

Zehner, S. P. and M. T. Tuley. Development and Validation of Multipath and Clutter Models for TAC ZINGER in Low Altitude Scenarios. Final Report of Project No. A-2194 for Headquarters USAF/SAGF under contract No. F49620-78-C-0121. Atlanta, GA: Engineering Experiment Station, Georgia Institute of Technology, March 1979.

Appendix A. Tilt Transformations

This appendix gives a derivation of the transformations of the terrain slopes from an X-Y coordinate system to a longitudinal-transverse (L-T) coordinate system. The coordinate systems are coplanar; the L and T axes are formed by rotating the X and Y axes through an angle B about the Z axis (Figure 14). The longitudinal tilt is the tilt component of the facet normal vector along the longitudinal axis. The transverse tilt is defined similarly.

Let $\hat{u}$ be the unit normal vector to the facet, with components u_x, u_y, and u_z in the X-Y-Z coordinate system, and components u_L, u_T, and u_Z in the L-T-Z system. Denoting the tilt components in the X, Y, L, and T directions as XTILT, YTILT, LTILT, and TTILT,

$$XTILT = \tan^{-1} (u_z/u_x) \tag{14}$$

$$YTILT = \tan^{-1} (u_z/u_y) \tag{15}$$

$$LTILT = \tan^{-1} (u_Z/u_L) \tag{16}$$

$$TTILT = \tan^{-1} (u_Z/u_T) \tag{17}$$

To express LTILT and TTILT in terms of XTILT and YTILT,

$$u_z = u_L \cos B \tan (XTILT) + u_L \sin B \tan (YTILT) \tag{18}$$

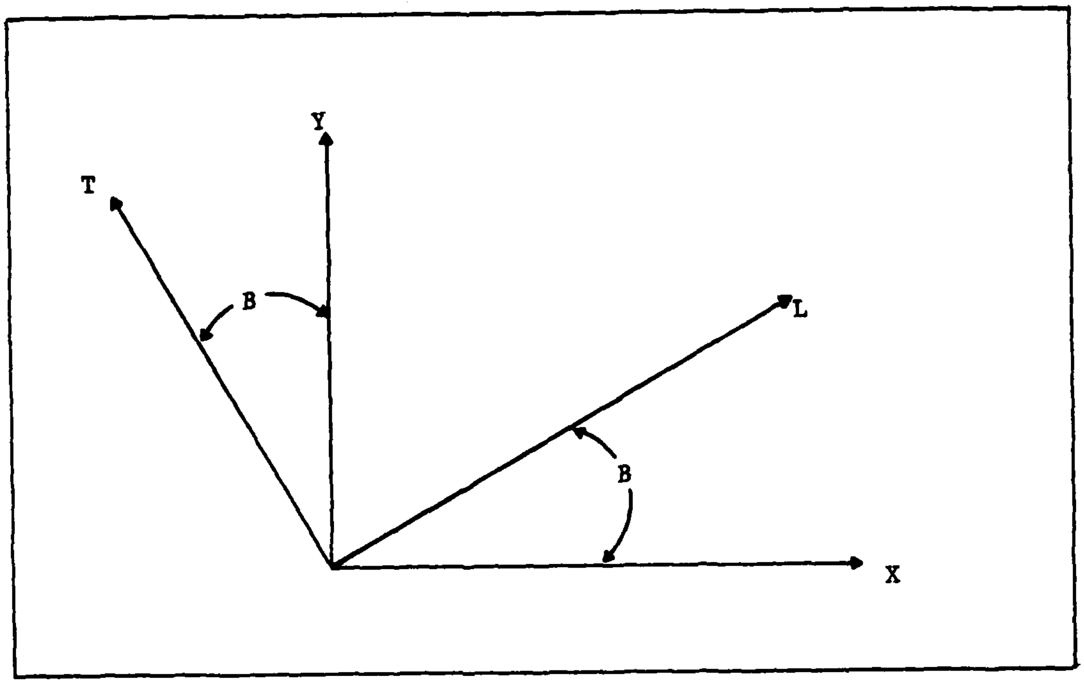

Figure 14. Rotation of X-Y Coordinate System Through Angle B to Form L-T System

where $u_L \cos B$ is the component of u_L in the X direction, and $u_L \sin B$ is the component of u_L in the Y direction. Thus, using equation 16

$$u_Z/u_L = \tan(\text{LTILT}) = \cos B \tan(\text{XTILT}) + \sin B \tan(\text{YTILT}) \qquad (19)$$

Using TTILT,

$$u_Z = (u_T)(-\sin B) \tan(\text{XTILT}) + u_T \cos B \tan(\text{YTILT}) \qquad (20)$$

where $(u_T)(-\sin B)$ is the component of u_T in the X direction, and $u_T \cos B$ is the component of u_T in the Y direction. Using equation 17,

$$u_Z/u_T = \tan(\text{TTILT}) = -\sin B \tan(\text{XTILT}) + \cos B \tan(\text{YTILT}) \qquad (21)$$

If XTILT, YTILT, LTILT, and TTILT are small, say less than 20°, the following approximation may be used:

$$\tan \theta \approx \theta \tag{22}$$

for θ an angle in radians.

Using equation 22 in equations 18 and 19 results in

$$\text{LTILT} \approx (\cos B)(\text{XTILT}) + (\sin B)(\text{YTILT}) \tag{23}$$

$$\text{TTILT} \approx (-\sin B)(\text{XTILT}) + (\cos B)(\text{YTILT}) \tag{24}$$

These are the equations used in FACET to make the coordinate transformations.

Appendix B. Computer Program

```
      PROGRAM GTTHES(INPUT=65,OUTPUT=65,TERAIN,CPARAM,DATA,TAPE5=INPUT,
     $ TAPE6=OUTPUT, TAPE30=CPARAM, TAPE31=TERAIN, TAPE7=DATA, INDATA,
     $ TAPE8=INDATA)

*********************************************************************
*                                                                   *
*     THIS IS A DRIVER PROGRAM FOR THE GEORGIA TECH CLUTTER AND     *
*     MULTIPATH MODEL DEVELOPED FOR AIR FORCE STUDIES AND ANALYSIS  *
*     FOR IMPLEMENTATION IN TAC ZINGER.  THIS PROGRAM WAS BUILT AS  *
*     PART OF THE RESEARCH FOR A MASTER'S DEGREE THESIS BY CAPTAIN  *
*     JAMES M. ARNETT, GOR 79-D, AIR FORCE INSTITUTE OF TECHNOLOGY. *
*                                                                   *
*     PROGRAM CURRENT AS OF 7 DEC 79.                               *
*                                                                   *
*********************************************************************
      COMMON/THESIS/FREQI,BWI,PWI,AI,BI,CI,DI,DEC2,ITHESIS
      COMMON/CTARA/WAVEHT,ITARA(5,40,40),TSCALE,IWET,ISNOW,ITTYPE
      COMMON/UTM/ITERSW,YNORTH,XEAST,PHI,ZMASK,XTILT,YTILT,RMS
C     THE FOLLOWING IS THE ORIGINAL UTM COMMON BLOCK.  IT INCLUDES
C     SEVERAL VARIABLES THAT ARE NEVER USED.
C     COMMON/UTM/ITERSW,YNORTH,XEAST,PHI,DRATE,IGRID(4,3),
C    $           RHMASK,XMASK,YMASK,ZMASK,IGRDR,ITERR(1717)

      COMPLEX DEC2,DEC2I
      REAL LTILT,LTILTI

C    **USE AVELERR AND AVAZERR FOR STRAIGHT AND LEVEL FLIGHT PATH, WITH
C    **DIMENSIONS AT LEAST AS LARGE AS NSTOP.
C    **DIMENSION AVELERR(500),AVAZERR(500)

      NAMELIST/INITIAL/IRUN,RANGE,ZT,SIGD,ZS,FREQI,BWI,PWI,SCVD,
     $ AI,BI,CI,DI,DEC2,RMS,LTILT,TTILT,ITERSW

C     NAMELIST "INITIAL" PARAMETERS
C        IRUN = RUN NUMBER
C        RANGE = GROUND DISTANCE FROM SITE TO TARGET (FEET)
C        SIGD = RADAR CROSS SECTION OF TARGET (SQUARE FEET IN DECIBELS)
C        ZT = HEIGHT OF TARGET ABOVE REFERENCE PLANE (FEET)
C        ZS = HEIGHT OF ANTENNA ABOVE REFERENCE PLANE (FEET)
C        FREQI = RADAR FREQUENCY (GHZ)
C        BWI = RADAR BEAMWIDTH (DEGREES)
C        PWI = RADAR PULSEWIDTH (NANOSECONDS)
C        SCVD = SUB-CLUTTER VISIBILITY (DB).  RADAR CAN DETECT MOVING
C               TARGET IN PRESENCE OF CLUTTER EVEN THOUGH CLUTTER ECHO
C               POWER IS (10**(SCVD/10)) TIMES THE TARGET ECHO POWER.
C        AI,BI,CI,DI = LOW ANGLE TERRAIN CLUTTER PARAMETERS (SEE
C                      DOCUMENTATION)
C        DEC2 = COMPLEX DIELECTRIC CONSTANT FOR DRY TERRAIN (SEE
C               DOCUMENTATION).
C        RMS = ROOT MEAN SQUARE SURFACE ROUGHNESS (CM)
```

```
C         LTILT = ANGLE BETWEEN Z AXIS AND COMPONENT OF FACET NORMAL IN
C                 L-Z PLANE (DEGREES)--LTILT IS POSITIVE IF THE COMPONENT
C                 OF THE FACET NORMAL ALONG THE L AXIS IS POSITIVE
C                 (L AXIS=LONGITUDINAL AXIS)
C         TTILT = ANGLE BETWEEN Z AXIS AND COMPONENT OF FACET NORMAL IN
C                 T-Z PLANE (DEGREES)--TTILT IS POSITIVE IF THE COMPONENT
C                 OF THE FACET NORMAL ALONG THE T AXIS IS POSITIVE
C                 (T AXIS=TRANSVERSE AXIS)
C                 NOTE:  THE L-T AXES ARE SIMPLY THE X-Y AXES ROTATED
C                 ABOUT THE Z AXIS SO THAT THE NEW X AXIS (L AXIS) IS
C                 PARALLEL  TO THE GROUND LINE (THE LINE IN THE X-Y PLANE
C                 THAT CONNECTS THE X,Y COORDINATES OF THE SITE AND THE
C                 TARGET.)
C         ITERSW = 0 IF LTILT EQ 0 AND TTILT EQ 0
C         ITERSW = 2 IF LTILT NE 0 OR TTILT NE 0


C      MAKE 32 RUNS, EACH WITH A DIFFERENT SET OF INPUT VALUES.
       DO 1000 IRUNNO=1,32

       READ(5,INITIAL)
       IF(EOF(5).NE.0) STOP "RAN OUT OF DATA"
       WRITE(6,INITIAL)
C      WRITE(8,INITIAL)

C      SAVE DATA FOR USE IN EACH REPLICATION.
       IRUNI=IRUN
       RANGEI=RANGE
       ZTI=ZT
       SIGDI=SIGD
       ZSI=ZS
       FREQII=FREQI
       BWII=BWI
       PWII=PWI
       SCVDI=SCVD
       AII=AI
       BII=BI
       CII=CI
       DII=DI
       DEC2I=DEC2
       RMSI=RMS
       LTILTI=LTILT
       TTILTI=TTILT

C     **USE CARDS LABELED WITH ** FOR STRAIGHT AND LEVEL FLIGHT PATH.
C     **FLY TARGET IN FROM A RANGE OF NSTOP*100 FEET AT A CONSTANT
C     **ALTITUDE.
C     **NSTOP=500
C     **DO 1000 IRUNNO=1,NSTOP
C     **RANGEI = NSTOP*100+100 - IRUNNO*100
C     **AZERTOT = 0.
C     **ELERTOT = 0.
```

```
C     USE SAME RANDOM NUMBER STREAM FOR EACH RUN.
      CALL RANSET(7)

C     MAKE NREP REPLICATIONS OF THE PROGRAM FOR EACH SET OF INPUT VALUES
      NREP=30
      DO 100 IREP=1,NREP

C     REINITIALIZE PARAMETERS IN NAMELIST "INITIAL".
      IRUN=IRUNI
      RANGE=RANGEI
      ZT=ZTI
      SIGD=SIGDI
      ZS=ZSI
      FREQI=FREQII
      BWI=BWII
      PWI=PWII
      SCVD=SCVDI
      AI=AII
      BI=BII
      CI=CII
      DI=DII
      DEC2=DEC2I
      RMS=RMSI
      LTILT=LTILTI
      TTILT=TTILTI

C     INITIALIZE INTERNAL PARAMETERS

      KEYSAM=0
      ITHESIS=1
C     ITHESIS=1 ==> THESIS RUN
C     ARBITRARILY SET AZT=5 AND ITTYPE=3 FOR THESIS RUN
      AZT=5.
      ITTYPE=3

C     ORIGIN OF SITE-CENTERED COORDINATE SYSTEM IS CO-LOCATED WITH
C     ORIGIN OF REFERENCE SYSTEM, AND THE AXES IN BOTH SYSTEMS ARE
C     PARALLEL.
      XEAST=0.
      XS=XEAST
      YNORTH=0.
      YS=YNORTH
      STHETA=0.
      SPHI=0.
      SPSI=0.

      ZMASK=0.
C     ZMASK = HEIGHT OF AN OBJECT BETWEEN SITE AND TARGET THAT PREVENTS
C             SITE FROM "SEEING" TERRAIN BELOW THAT HEIGHT.

C     CONVERT DEGREES TO RADIANS
      DTOR=4.*ATAN(1.)/180
      AZT=AZT*DTOR
      SPSI=SPSI*DTOR
```

```
      STHETA=STHETA*DTOR
      SPHI=SPHI*DTOR
      LTILT=LTILT*DTOR
      TTILT=TTILT*DTOR

      PHI=SPHI


C     CALCULATE XTILT AND YTILT GIVEN LTILT, TTILT AND AZT.

C     XTILT = TILT OF FACET NORMAL IN X DIRECTION (RADIANS)--
C             POSITIVE=FACET EDGE CLOSER TO ORIGIN IS HIGHER THAN
C             FARTHER EDGE.
C     YTILT = TILT OF FACET NORMAL IN Y DIRECTION (RADIANS)--
C             POSITIVE=FACET EDGE CLOSER TO ORIGIN IS HIGHER THAN
C             FARTHER EDGE.
      TANLT=TAN(LTILT)
      TANTT=TAN(TTILT)
      SINAZT=SIN(AZT)
      COSAZT=COS(AZT)
      TANAZT=TAN(AZT)
      TANYT=(TANTT+TANAZT*TANLT)/(SINAZT*TANAZT+COSAZT)
      TANXT=(TANLT-SINAZT*TANYT)/COSAZT
      XTILT=ATAN(TANXT)
      YTILT=ATAN(TANYT)
C     EXPRESS XTILT AND YTILT IN MILLIRADIANS
      XTILT=XTILT*1000
      YTILT=YTILT*1000

C     CALCULATE COORDINATES OF TARGET IN REFERENCE SYSTEM
      XT=XS+RANGE*COS(AZT)
      YT=YS+RANGE*SIN(AZT)

C     CALCULATE THE ELEVATION ANGLE FROM THE ANTENNA TO THE TARGET.
      ZDIF=ZT-ZS
      ELT=ATAN3(ZDIF,RANGE)

C     CALCULATE ANTENNA'S ELEVATION AND AZIMUTH ANGLES.  I.E., WHERE THE
C     ANTENNA IS POINTING.
C     ASSUME PERFECT TRACKING
      ELS=ELT
      AZS=AZT

      CALL MULTIN(KEYSAM,SCVD)

      CALL MULTIP(XS,YS,ZS,XT,YT,ZT,ZMASK,SIGD,SCVD,AZS,ELS,AZT,ELT,
     $ AZERR,ELERR)

C     CORRECT ELERR IF ELEVATION ANGLE IS < 1 DEGREE.  SEE FINAL REPORT
C     OF GEORGIA TECH'S DEVELOPMENT OF THE MODEL, PAGE 67.
      IF(ELS/DTOR.LT.1.0) ELERR=ELERR+(1-ELS/DTOR)*DTOR

C    **FOR FLIGHT PATH, SUM ERRORS AT EACH INCREMENT TO GET AVERAGE.
C    **AZERTOT=AZERTOT+AZERR
```

```
C    **ELERTOT=ELERTOT+ELERR

      WRITE(6,10) IRUNNO,IREP,AZERR,ELERR,SIGD
 10   FORMAT(" RUN ",I3,"   REP ",I3,"   AZERR = ",E21.15,"   ELERR = ",
     $ E21.15,"   SIGD = ",E21.15)

C     DEFINE HIGH AND LOW VALUES OF VARIABLES TO BE +1. AND -1. FOR DATA
C     ANALYSIS.  DA1=RANGE,DA2=ZT,...,DA16=TTILT.
      DA1 =1.
      DA2 =1.
      DA3 =1.
      DA4 =1.
      DA5 =1.
      DA6 =1.
      DA7 =1.
      DA8 =1.
      DA9 =1.
      DA10=1.
      DA11=1.
      DA12=1.
      DA13=1.
      DA14=1.
      DA15=1.
      DA16=1.
      IF(RANGEI.EQ.3000.)       DA1 =-1.
      IF(ZTI.EQ.100.)           DA2 =-1.
      IF(SIGDI.EQ.-9.68)        DA3 =-1.
      IF(ZSI.EQ.6.)             DA4 =-1.
      IF(FREQII.EQ.9.)          DA5 =-1.
      IF(BWII.EQ..8)            DA6 =-1.
      IF(PWII.EQ.100.)          DA7 =-1.
      IF(SCVDI.EQ.1.)           DA8 =-1.
      IF(AII.EQ..0045)          DA9 =-1.
      IF(BII.EQ..83)            DA10=-1.
      IF(CII.EQ..0013)          DA11=-1.
      IF(DII.EQ.0.)             DA12=-1.
      IF(DEC2I.EQ.(2.,0.))      DA13=-1.
      IF(RMSI.EQ.10.)           DA14=-1.
      IF(LTILTI.EQ.0.)          DA15=-1.
      IF(TTILTI.EQ.0.)          DA16=-1.

C     WRITE OUTPUT AND INPUT FOR EACH REPLICATION TO TAPE7 FOR DATA
C     ANALYSIS.
      WRITE(7,11)AZERR,ELERR,SIGD,DA1,DA2,DA3,DA4,DA5,DA6,DA7,DA8,DA9,
     $ DA10,DA11,DA12,DA13,DA14,DA15,DA16
 11   FORMAT(3E13.6,10F4.0/6F4.0)

 100  CONTINUE

C    **AVAZERR(IRUNNO)=AZERTOT/NREP
C    **AVELERR(IRUNNO)=ELERTOT/NREP
C    **WRITE(6,13)RANGE,AVAZERR(IRUNNO),AVELERR(IRUNNO)
C13 **FORMAT("   RANGE = ",F7.0,"     AVAZERR = ",1PE13.6,
C  **$ "     AVELERR = ",1PE13.6)
```

```
 1000 CONTINUE

C    **WRITE(7,12)((AVAZERR(J),AVELERR(J)),J=1,NSTOP)
C12 **FORMAT(2E13.6)

      STOP
      END
```

```
      SUBROUTINE MULTIN(KEYSAM,SCVD)
C
C          REVISION 6.03
C     MOD BY JJOHNSON  7 FEB 79 TO BY-PASS MULTIN FOR NO TERRAIN
C     MOD BY JJOHNSON 23 FEB 79 TO INITIALIZE FREQ, BWE, BWA, AND PW
C     MOD BY SPSTUK 26 FEB 79 TO READ DEFAULT TERRAIN
C
      COMMON/CTARA/WAVEHT,ITARA(5,40,40),TSCALE,IWET,ISNOW,ITTYPE
      DIMENSION ETA(80),THETA1(80),RHOD(80)
      DIMENSION A(9),B(9),C(9),D(9),DC(9)
      COMPLEX RHO,DEC,ZERO
C     THE FOLLOWING PARAMETERS WERE IN THE ORIGINAL GATECH PROGRAM, BUT
C     ARE NOT NEEDED IN MULTIN.
C     DIMENSION IFACET(5,80),POS(4,80)
C     COMPLEX RHOP,FAC2,FAC3,ZTRI,FF,FE,FA
      COMMON/MULCLUT/ RHO(80),DEC(9,2),ETA,THETA1,RHOD,A,B,C,D,DC,ALVB
     $,PI,C0,C01,C02,C1,C2,C3,C31,C4,C5,C6,TS,RK,BWE,BWE1
     $,BWA,BWA1,PE,PA,FSCVD,IRAN,LOOP,IFCL,NFACET,RSIGC,FG,ZERO
     $,WL,FREQ,PW,R,RTC,ISPSW
      COMMON/THESIS/FREQI,BWI,PWI,AI,BI,CI,DI,DEC2,ITHESIS
      COMPLEX DEC2
      COMMON/UTM/ITERSW,YNORTH,XEAST,PHI,ZMASK,XTILT,YTILT,RMS
C     THE FOLLOWING IS THE ORIGINAL UTM COMMON BLOCK.  IT INCLUDES
C     SEVERAL VARIABLES THAT ARE NEVER USED.
C     COMMON/UTM/ITERSW,YNORTH,XEAST,PHI,DRATE,IGRID(4,3),
C    $           RHMASK,XMASK,YMASK,ZMASK,IGRDR,ITERR(1717)

      ZERO=CMPLX(0.,0.)
      DO 983 I=1,80
      THETA1(I)=0.
      ETA(I)=0.
      RHO(I)=ZERO
983   RHOD(I)=0.

C     INITIALIZE FREQUENCY, BEAMWIDTH, AND PULSEWIDTH

      IF(ITHESIS.EQ.1) GO TO 199
C     ITHESIS=1 ==> THESIS RUN

C     FREQUENCY IN GHZ
      FREQ=5.28
C     BEAMWIDTHS IN DEGREES.
      BWA=1.1
      BWE=1.1
C     PULSEWIDTH IN NANOSECONDS
      PW=250
      GO TO 200
199   FREQ=FREQI
      BWE=BWI
      BWA=BWI
      PW=PWI
200   WL=1/FREQ
      PI=4.*ATAN(1.)
      C0=SQRT(2.)*PI*2./(WL*30.48)
```

```
      C01=4.*SQRT(PI)
      C02=PI/2.
      IF(ITERSW.NE.1) GO TO 501
      READ(31,876) TSCALE
876   FORMAT(2F12.0,2I6)
      GO TO 502
C     DEFAULT VALUES
501   TSCALE=1650.
      WAVEHT=1.
      ISNOW=0
      IWET=2
502   TS=1./TSCALE
      RK=PI*TSCALE*2./WL
C     BEAMWIDTH IN ELEVATION (DEGREES TO RADIANS)
      BWE=BWE*PI/180.
      BWE1=.7336*BWE
C     BEAMWIDTH IN AZIMUTH (DEGREES TO RADIANS)
      BWA=BWA*PI/180.
      BWA1=.7336*BWA
      PE=.001*FS(.001,BWE1)/FD(.001,BWE1)
      PA=.001*FS(.001,BWA1)/FD(.001,BWA1)
      C1=PI/(2.*RK)
      C2=RK/PI
      RTC=0.
      R=0.
      ISPSW=1
C     C*TAU/2 IN TSCALE UNITS (C = APPROXIMATELY 1 FT/NANOSECOND)
C     TAU=PULSE WIDTH
      C3=PW*.5*TS
      C31=C3/SQRT(2.)
      C4=(5.7887*WAVEHT/.1156)**4
      FAC1=1.7/(WL+.05)**.4
      C5=10.47*WAVEHT**0.4
      C5=(C5/(1.+C5/30.))**FAC1
      C5=4.226E-7*C5*(WAVEHT+.05)**(-.24)*WL*(WL+.05)**.25
      C6=2.3779/BWE
      FSCVD=EXP(-SCVD/8.686)
      IRAN=1+TSCALE/150.
      LOOP=-1
      IFCL=0

C     IF THIS IS A THESIS RUN, SKIP TO THESIS INPUT.
      IF(ITHESIS.EQ.1) GO TO 1000

C     GROUND CORRELATION
      DO 984 I=1,9
984   DC(I)=TSCALE*.3048*.37
      DC(1)=30.
C     READ CLUTTER $ MULTIPATH MODEL PARAMETERS FROM DATA FILE
      DO 660 I=1,9
      READ(30,661) A(I),B(I),C(I),D(I)
661   FORMAT(4F10.0)
660   CONTINUE
      DO 662 I=1,9
```

```
      READ(30,661) DEC(I,1),DEC(I,2)
662   CONTINUE


C     DEFINE TERRAIN

C     NOTE:  TERRAIN DATA READ INTO ITARA AS ITARA(K,J,I), NOT AS
C            ITARA(K,I,J).

C        J           =Y COORDINATE OF FACET IN TSCALE UNITS
C        I           =X COORDINATE OF FACET IN TSCALE UNITS
C        ITARA(1,J,I)=AVERAGE CENTER HEIGHT OF FACET (CM)
C        ITARA(2,J,I)=TILT IN X DIRECTION (MILLIRADIANS)
C        ITARA(3,J,I)=TILT IN Y DIRECTION (MILLIRADIANS)
C        ITARA(4,J,I)=ROOT MEAN SQUARE SURFACE ROUGHNESS (CM)
C        ITARA(5,J,I)=TERRAIN TYPE

      IMAX=40
      JMAX=40

      IF(ITERSW.EQ.1) GO TO 503

C     DEFAULT TERRAIN VALUES
      DO 504 J=1,JMAX
      DO 504 I=1,IMAX
      ITARA(1,J,I)=0
      ITARA(2,J,I)=0
      ITARA(3,J,I)=0
      ITARA(4,J,I)=10
      ITARA(5,J,I)=ITTYPE
504   CONTINUE
      GO TO 880

C     READ TERRAIN FROM FILE 31.
503   DO 700 J=1,JMAX
      DO 700 I=1,IMAX
      DO 700 K=1,5
      ITARA(K,J,I)=0
700   CONTINUE
      DO777 JJ=1,1600
      READ(31,878)          INEW,JNEW,(ITARA(I,JNEW,INEW),I=1,5)
      IF(EOF(31).NE.0.) GO TO 880
  878 FORMAT(7I6)
777   CONTINUE


C     DEFINE PARAMETERS FOR THESIS RUN.
1000  DO 1001 I=1,9
      DC(I)=TSCALE*.3048*.37
      DC(1)=30.
      A(I)=AI
      B(I)=BI
      C(I)=CI
      D(I)=DI
```

```
      DEC(I,1)=CMPLX(0.,0.)
1001  DEC(I,2)=DEC2

C     DEFINE TERRAIN FOR THESIS RUN.
      IMAX=40
      JMAX=40
      DO 1002 J=1,JMAX
      DO 1002 I=1,IMAX
      ITARA(1,J,I)=0
      ITARA(2,J,I)=XTILT
      ITARA(3,J,I)=YTILT
      ITARA(4,J,I)=RMS
1002  ITARA(5,J,I)=ITTYPE

880   CONTINUE
C     WRITE(6,1003)
1003  FORMAT(/" TERRAIN TYPE",5X,"A",9X,"B",9X,"C",9X,"D",12X,"DEC1",
     $  16X,"DEC2"//)
C     WRITE(6,1004) ((I,A(I),B(I),C(I),D(I),DEC(I,1),DEC(I,2)),I=1,9)
1004  FORMAT(6X,I1,6X,8F10.5)
C880  WRITE(6,640)
 640  FORMAT(/" END OF INITIALIZE SEGMENT IN MULTIPATH."//)
      RETURN
      END
```

```
      SUBROUTINE MULTIP(XS,YS,ZS,XT,YT,ZT,ZMASK,SIGD,SCVD
     1,AZS,ELS,AZT,ELT,AZERR,ELERR)

C          REVISION 6.02

      DIMENSION IFACET(5,80),POS(4,80),ETA(80),THETA1(80),RHOD(80)
      DIMENSION A(9),B(9),C(9),D(9),DC(9)
      COMPLEX RHO,RHOP,DEC,FAC2,FAC3,ZTRI,FF,FE,FA,ZERO
     $,RHOSUM
      LOGICAL IDSPEC
      COMMON/CTARA/WAVEHT,ITARA(5,40,40),TSCALE,IWET,ISNOW,ITTYPE
      COMMON/MULCLUT/ RHO(80),DEC(9,2),ETA,THETA1,RHOD,A,B,C,D,DC,ALVB
     $,PI,C0,C01,C02,C1,C2,C3,C31,C4,C5,C6,TS,RK,BWE,BWE1
     $,BWA,BWA1,PE,PA,FSCVD,IRAN,LOOP,IFCL,NFACET,RSIGC,FG,ZERO
     $,WL,FREQ,PW,R,RTC,ISPSW
      COMMON/UTM/ITERSW,YNORTH,XEAST,PHI,DUMMY,XTILT,YTILT,RMS
C     THE FOLLOWING IS THE ORIGINAL UTM COMMON BLOCK.  IT INCLUDES
C     SEVERAL VARIABLES THAT ARE NEVER USED.
C     COMMON/UTM/ITERSW,YNORTH,XEAST,PHI,DRATE,IGRID(4,3),
C    $           RHMASK,XMASK,YMASK,DUMMY,IGRDR,ITERR(1717)

C          *******FACET PARAMETERS*******
C     XS, YS, ZS ARE THE COORDINATES OF THE START POINT
C     XT, YT, ZT ARE THE COORDINATES OF THE TERMINUS
C     IFACET(1,I) = TRANSVERSE TILT (MR)
C     IFACET(2,I) = LONGITUDINAL TILT (MR)
C     IFACET(3,I) = RMS SURFACE ROUGHNESS (CM)
C     IFACET(4,I) = TERRAIN TYPE
C     IFACET(5,I) = 1 IF THE FACET IS NOT SHADOWED, ELSE IT IS SHADOWED
C     POS(1,I) = FACET HEIGHT (TSCALE UNITS)
C     POS(2,I) = GROUND LINE SEGMENT LENGTH (TSCALE UNITS)
C     POS(3,I) = AVERAGE ELEVATION ANGLE FROM START POINT (RADIANS)
C     POS(4,I) = AVERAGE ELEVATION ANGLE FROM TERMINUS (RADIANS)
C     NFACET RETURNS THE # FACETS FOUND
C     ISHAD=0 => CLUTTER CALL TO FACET--DO NOT CHECK FOR SHADOWING
C           1 => MULTIPATH CALL TO FACET--CHECK FOR SHADOWING
C     NOSHAD=0 => DELETE ALL SHADOWING TESTS (PARAMETER ADDED FOR THESIS)
C            1 => CHECK FOR SHADOWING

C        AZERR=0.
C        ELERR=0.
C        GO TO 9999
C 111 CONTINUE
C      WRITE(6,191) XS,YS,ZS,XT,YT,ZT
C      WRITE(6,191) AZS,ELS,AZT,ELT,ZMASK
C191   FORMAT(5F20.6)
C*    COMPUTE TRACK ERRORS AND WRITE TO FILE(EGLIN)
C     WELER=(ELT-ELS)*1000.
C     WAZER=(AZT-AZS)*1000.
C     RTC=RTC+1.
C     WRER=0.
C     WR=TSCALE*R
C*    REMOVES ZEROES FROM OUTPUT TAPE
C     IF(AZT.EQ.0) GO TO 663
```

```
C      WRITE(3,664) RTC,WR,AZT,ELT,WRER,WAZER,WELER
C664   FORMAT(F18.0,F14.0,2F10.4,F10.0,2F10.4)
C*    PRECEEDING LINES FOR OUTPUT TAPE


C     PRINT*,"MULTIP ENTRY:  ",XS,YS,ZS,XT,YT,ZT,ZMASK,SIGD,SCVD,AZS,
C    $ ELS,AZT,ELT
C     COMPUTE HEIGHTS AND RANGES
663    POSS=ZS*TS
      POST=ZT*TS
C     IF ZMASK=0. AND TERRAIN IS NOT FLAT, SET POSZ TO HEIGHT OF TERRAIN
C     DIRECTLY BENEATH THE TARGET.
      POSZ=ZMASK*TS
      G=SQRT((XT-XS)**2+(YT-YS)**2)*TS
C     R=SLANT RANGE FROM SITE TO TARGET.
      R=SQRT(G*G+(POST-POSS)**2)
C      GO TO 998
C     IF TARGET IS MORE THAN 20 BEAMWIDTHS ABOVE THE TERRAIN, MULTIPATH
C     AND CLUTTER EFFECTS ARE ASSUMED NEGLIGIBLE.  GO TO FREESPACE
C     CALCULATIONS.
      IF(POST-POSZ.GT.20.*BWE*G)GOTO 998
      FSIGD=EXP(SIGD/8.686)
C     IN ORIGINAL PROGRAM, ONLY CHECKED MULTIPATH AND CLUTTER EVERY
C     IRAN CALLS TO MULTIP.
C     LOOP=MOD(LOOP+1,IRAN)
C     LOOP = 0 IMPLIES CHECK MULTIPATH AND CLUTTER AT EVERY DT.
C     IF(LOOP.NE.0)GOTO 999


C           * * * * * * MULTIPATH SECTION  * * * * * * * * * * * * *

      CALL FACET(XS,YS,ZS,XT,YT,ZT,IFACET,POS,NFACET,1,0)
      OLPOS=0.
      OLDBETA=-1.
      IDSPEC=.FALSE.
      G1=0.
C      SRL=0.

      DO 997 I=1,NFACET
      IF4=IFACET(4,I)
      IF(IF4.GT.0) GO TO 301
C     WRITE(6,300)
300   FORMAT(5X,"  OFF TERRAIN  ")
      GO TO 997
301   IF(ISNOW.EQ.1) IF4=9
      RHOD(I)=0.
      RHO(I)=ZERO
C     G1=GROUND LINE DISTANCE FROM INITIAL TARGET POSITION TO CENTER OF
C     GROUND LINE SEGMENT IN CURRENT FACET.
      G1=G1+(POS(2,I)+OLPOS)*.5
      OLPOS=POS(2,I)
C      IF FACET IS SHADOWED GO TO 997.
      IF(IFACET(5,I).NE.1) GOTO 997
      G2=G-G1
```

```
      H1=POSS-POS(1,I)
      IF(ELS+H1/G1.GT.20.*BWE)GOTO 997
      R1=SQRT(G1*G1+H1*H1)
      H2=POST-POS(1,I)
      R2=SQRT(G2*G2+H2*H2)
      DEL0=R1+R2-R
      IF(DEL0.GE.C3)GOTO 997
      BETA0=0.02*IFACET(3,I)/DC(IF4)
C     THETA1 = AVERAGE ELEVATION ANGLE FROM START POINT
      THETA1(I)=POS(3,I)
C     ALPU = LONGITUDINAL TILT OF FACET
      ALPU=IFACET(2,I)*.001
C     PSI1 = INCIDENCE ANGLE
      PSI1=THETA1(I)-ALPU
C     IF INCIDENCE ANGLE LESS THAN ZERO, THE FACET IS SELF-SHADOWED, SO
C     GO TO NEXT FACET
      IF(PSI1.LE.0.) GO TO 997
      BETA=(-THETA1(I)+POS(4,I))*.5+ALPU
      ABETA=ABS(BETA)
      BETAM=C1*SQRT(1.+C2*DEL0)/(R1*PSI1)

      IF(OLDBETA.LT.0.0.AND.BETA.GT.0.0) ABETA=0.0
C     THIS FINDS THE FACET WHERE THE INCIDENCE AND REFLECTIONS ANGLES
C     ARE CLOSEST TO BEING EQUAL, WITH BETA GT 0.0.  THIS IS THE FACET
C     WITH THE STRONGEST SPECULAR REFLECTION.  FOR DEFAULT TERRAIN, IT
C     IS ASSUMED THAT ALL OF THE SPECULAR REFLECTION COMES FROM THIS
C     FACET.  NOTE:  RMS CANNOT EQUAL ZERO, OR BETA0 WILL ALSO BE ZERO,
C     AND WILL CAUSE A DIVISION BY ZERO LATER IN THE CODE.

      OLDBETA=BETA
      IF(ABS(BETA).GT.BETA0.AND.ABS(BETA).GT.BETAM) GO TO 997
      RHOS2=0.
      RHOS1=0.
      FAC1=C0*IFACET(3,I)
      IF(IF4.EQ.1) FAC1=C0*7.6*WAVEHT
      FAC4=FAC1*PSI1
      IF(FAC4.LT.7.) RHOS1=EXP(-1.*FAC4**2)
      FAC4=FAC1*(POS(4,I)+ALPU)
      IF(FAC4.LT.7.) RHOS2=EXP(-1.*FAC4**2)
      REALP=REAL(DEC(IF4,IWET))
      IF(REALP.NE.0.)GOTO 995
      RHOP=CMPLX(-AIMAG(DEC(IF4,IWET)),0.)
      GOTO 996
995   FAC2=CSQRT(DEC(IF4,IWET)+PSI1**2-1.)
      FAC3=DEC(IF4,IWET)*PSI1
      RHOP=(FAC3-FAC2)/(FAC3+FAC2)
996   ETA(I)=-.001*IFACET(1,I)*(ELT+THETA1(I))
C     TEST FOR DIFFUSE MULTIPATH.
      IF(ABS(BETA).GT.BETA0) GO TO 994
      RHOS1S=0.
      RHOS2S=0.
      IF(RHOS1.GT.1.0E-10) RHOS1S=RHOS1*RHOS1
      IF(RHOS2.GT.1.0E-10) RHOS2S=RHOS2*RHOS2
      FAC1=SQRT((1.-RHOS1S)*(1.-RHOS2S))
```

```
      DR1=POS(2,I)/COS(PSI1)
C     INSURE ARGUMENT OF EXPONENT IS NOT OUT OF LIMITS
      IF(((BETA/BETAO)**2).GT.675.) GO TO 502
      RHOD2=R*FAC1*EXP(-1.*(BETA/BETAO)**2)*(THETA1(I)+POS(4,I))*DR1*
     ,(REAL(RHOP)**2+AIMAG(RHOP)**2)/(R1*R2*CO1*BETAO)
      GO TO 503
502   RHOD2=0.0
503   RHOD(I)=SQRT(RHOD2)

C     TEST FOR SPECULAR MULTIPATH.
      IF(ABS(BETA).GT.BETAM) GO TO 997
C     EQ 2.2-4
994   RL=G*SQRT(1.+RK*DELO/PI)/(1.+RK*(H1+H2)**2/(2.*PI*G))

C     THE FOLLOWING LINES, UP TO 504, ARE INCLUDED TO GIVE BETTER
C     SPECULAR RESULTS FOR DEFAULT TERRAIN.
      WAITF=POS(2,I)/RL
C     WAITF=WEIGHTING FACTOR
      IF(ITERSW.NE.0) GO TO 504
      WAITF=0.
      IF(IDSPEC) GO TO 504
      IF(ABETA.NE.0.) GO TO 504
      IDSPEC=.TRUE.
      WAITF=1.
504   RHO(I)=CMPLX(COS(RK*DELO),-SIN(RK*DELO))*RHOP*RHOS1*WAITF
997   CONTINUE

C     DETERMINE NUMBER OF LAST FACET THAT CONTRIBUTES TO MULTIPATH.  IF
C     NFACET=0 AFTER THESE CALCULATIONS, NO MULTIPATH.
      NSTOP=NFACET+1

      DO 982 I=1,NSTOP
      NN=NFACET-I+1
      IF(NN.EQ.0)GOTO 981
      IF(RHO(NN).NE.ZERO .OR. RHOD(NN).NE.0.)GOTO 981
982   CONTINUE

981   NFACET=NN

C           * * * * * * END OF MULTIPATH SECTION * * * * * * * * * *

C?          USE GO TO 999  FOR BEACON
C       GO TO 999


C           * * * * * * CLUTTER SECTION  * * * * * * * * * * * * * *

      DISTL=R*.5*BWA*TSCALE
      FAC4=COS(AZS)
      FAC1=SIN(AZS)
      FACC=DISTL*FAC4
      FACS=DISTL*FAC1
      FACGC=FAC4*G*TSCALE
      FACGS=FAC1*G*TSCALE
```

```
      XA=XS+FACGC-FACS
      XB=XS+FACGC+FACS
      YA=YS+FACGS+FACC
      YB=YS+FACGS-FACC
      CALL FACET(XA,YA,0.,XB,YB,0.,IFACET,POS,NFAC,0,0)
      SIGOB=0.
      ALVB=0.
      THET1B=0.
      IFCL=1

      DO 993 I=1,NFAC
      IF(POS(1,I).LT.POSZ)GOTO 993
C     THETA = LINE OF SIGHT ANGLE FROM ANTENNA TO CLUTTER FACET.
      THETA=(POSS-POS(1,I))/G
C     GAMMA = INCIDENCE ANGLE OF BEAM TO CLUTTER FACET.
      GAMMA=THETA-IFACET(1,I)*.001
C     IF INCIDENCE ANGLE IS < 0, NO CLUTTER FROM THIS FACET.
      IF(GAMMA.LE.0.)GOTO 993
      IF4=IFACET(4,I)
      IF(ISNOW.EQ.1) IF4=9
      IF(IF4.EQ.1)GOTO 992
C     EQ2.3-5
      SIGO=A(IF4)*(GAMMA+C(IF4))**B(IF4)*
     1EXP(-D(IF4)/(1.+.03*IFACET(3,I)))
      IF(IF4.NE.9) GO TO 232
      IF(IWET.EQ.1) GO TO 991
      SIGO=SIGO*.316
      GO TO 991
232   IF(IWET.EQ.1)SIGO=SIGO*3.16
      GOTO 991

C     CLUTTER PATCH CALCULATIONS FOR SEA.
992   FAC1=C4*GAMMA**4
      FAC1=FAC1/(1.+FAC1)
      SIGO=C5*GAMMA**.7*FAC1
991   SIGOB=SIGOB+SIGO*POS(2,I)
C     SIGOB = SIGMA SUB ZERO BAR = AVG CLUTTER CROSS SECTION FOR THE
C     ENTIRE CLUTTER CELL.
      THET1B=THET1B+THETA*POS(2,I)
      ALVB=ALVB+0.001*IFACET(2,I)*POS(2,I)
993   CONTINUE

C     IF SIGOB VERY SMALL, NO CLUTTER.
      IF(SIGOB.LT.1.E-7*TS) GOTO 987
      FAC1=.5*TSCALE/DISTL
      THET1B=THET1B*FAC1
      ALVB=ALVB*FAC1
      FAC1=C6*ABS(THET1B+ELS)
      IF(FAC1.GT.2.)GOTO 990
      IF(ABS(FAC1-1.).LT..001)GOTO 989
      FG=COS(CO2*FAC1)/(1.-FAC1*FAC1)
      GOTO 988
990   FG=-1./(1.-FAC1*FAC1)
      GOTO 988
```

```
989    FG=1.4855-.7*FAC1
988    RSIGC=SQRT(SIGOB*C31)*FG*FG*FSCVD*TSCALE
       GO TO 999
987    IFCL=0

C            * * * * * * * END OF CLUTTER SECTION * * * * * * * * * * * * * *


999    FF=FS(ELT-ELS,BWE1)
       FE=FD(ELT-ELS,BWE1)
       TAI=AZT-AZS
       FA=FD(TAI,BWA1)*FF
       IF(NFACET.EQ.0)GOTO 980

       DO 985 I=1,NFACET
       TEI=-(ELS+THETA1(I))
       ZTRI=RHO(I)
       IF(RHOD(I).GT.0.)ZTRI=RICE(RHO(I),RHOD(I))
       FAC2=FS(TEI,BWE1)*ZTRI
       FF=FF+FAC2
       FE=FE+FD(TEI,BWE1)*ZTRI
       FA=FA+FD(TAI+ETA(I),BWA1)*FAC2
985    CONTINUE

980    SIGD=8.686*ALOG(REAL(FF)**2+AIMAG(FF)**2)+SIGD

C      IFCL=0 IMPLIES NO CLUTTER.
       IF(IFCL.EQ.0) GO TO 986

C      CALCULATE ERROR DUE TO BOTH MULTIPATH AND CLUTTER.
       ZTRI=RICE(ZERO,RSIGC)
       FAC2=FF*FSIGD
       FAC3=FF*FAC2+ZTRI
       AZERR=PA*REAL((FA*FAC2+FD(ALVB*(THET1B+ELS),BWA1)*
      1ZTRI/FG)/FAC3)
       ELERR=PE*REAL((FE*FAC2+FD(-THET1B-ELS,BWE1)*ZTRI/FG)/FAC3)
       GO TO 875

C      CALCULATE ERROR BASED ON MULTIPATH ONLY.
986    AZERR=PA*REAL(FA/FF)
       ELERR=PE*REAL(FE/FF)
       GO TO 875

C      CALCULATE ERROR IN FREE SPACE (NO MULTIPATH OR CLUTTER.)
998    FAC1=AZT-AZS
       FAC4=ELT-ELS
       AZERR=PA*FD(FAC1,BWA1)/FS(FAC1,BWA1)
       ELERR=PE*FD(FAC4,BWE1)/FS(FAC4,BWE1)
       LOOP=-1
       GO TO 875

875    RETURN
       END
```

```
      SUBROUTINE FACET(XSF,YSF,ZSF,XTF,YTF,ZTF,IFACET,POS,NFACET,ISHAD,
     $ NOSHAD)

           REVISION 5.03

      DIMENSION IFACET (5,80),POS(4,80),Z1(80),Z2(80)
      DIMENSION YNC(2,2),POINT(2,2),B(2),E(2),SYNE(2),DEL(2),ADEL(2)
      COMMON/CTARA/WAVEHT,ITARA(5,40,40),TSCALE,IWET,ISNOW,ITTYPE
C     XS, YS, ZS ARE THE COORDINATES OF THE START POINT
C     XT, YT, ZT ARE THE COORDINATES OF THE TERMINUS
C     IFACET(1,I) = TRANSVERSE TILT (MR)
C     IFACET(2,I) = LONGITUDINAL TILT (MR)
C     IFACET(3,I) = RMS SURFACE ROUGHNESS (CM)
C     IFACET(4,I) = TERRAIN TYPE
C     IFACET(5,I) = 1 IF THE FACET IS NOT SHADOWED, ELSE IT IS SHADOWED
C     POS(1,I) = FACET HEIGHT (TSCALE UNITS)
C     POS(2,I) = GROUND LINE SEGMENT LENGTH (TSCALE UNITS)
C     POS(3,I) = AVERAGE ELEVATION ANGLE FROM START POINT (RADIANS)
C     POS(4,I) = AVERAGE ELEVATION ANGLE FROM TERMINUS (RADIANS)
C     NFACET RETURNS THE # FACETS FOUND
C     ISHAD=0 => CLUTTER CALL TO FACET--DO NOT CHECK FOR SHADOWING
C            1 => MULTIPATH CALL TO FACET--CHECK FOR SHADOWING
C     NOSHAD=0 => DELETE ALL SHADOWING TESTS (PARAMETER ADDED FOR THESIS)
C             1 => CHECK FOR SHADOWING
C     ****************FACET FINDING DEFINITIONS*******************
C     IND=1 POINTS TO THE NEW POINT WITH X AN INTEGER
C     IND=2   "      "  "   "    "      "  Y  "   "
C     YNC(IND,1) HOLDS XINC, IND=1,2
C     YNC(IND,2) HOLDS YINC, IND=1,2
C     DEL(I)   HOLD XT-XS AND YT-YS
C     SYNE(I)   HOLD THE SIGNS OF DEL(I)
C     ADEL(I)   HOLD THE ABSOLUTE VALUES OF DEL(I)
C     ************************************************************
C     BEGIN GENERAL INITIALIZATION
C     CONVERT FT TO TSCALE UNITS (TSCALE IS IN FEET)
      CVFT=1./TSCALE
C     CONVERT CM TO TSCALE UNITS
      CVCM=CVFT/(12.*2.54)
      XS=XSF*CVFT
      YS=YSF*CVFT
      ZS=ZSF*CVFT
      XT=XTF*CVFT
      YT=YTF*CVFT
      ZT=ZTF*CVFT

C     CVTXY AND THE CONVERSIONS FROM METERS ARE NEEDED ONLY IF XTF, YTF,
C     XSF, AND YSF ARE DEFINED IN FEET USING A SITE-CENTERED COORDINATE
C     SYSTEM IN WHICH THE SITE IS NOT CO-LOCATED WITH THE ORIGIN OF A
C     REFERENCE SYSTEM DEFINED IN METERS, OR IN WHICH THE COORDINATE
C     AXES OF THE SITE-CENTERED SYSTEM ARE NOT PARALLEL TO THE REFERENCE
C     SYSTEM'S.
C        CVM=3.28/TSCALE
C        CALL CVTXY(XTF,YTF,XT,YT)
C        XT=XT*CVM
```

```
C        YT=YT*CVM
C        CALL CVTXY(XSF,YSF,XS,YS)
C        XS=XS*CVM
C        YS=YS*CVM

      RANGE=SQRT((XS-XT)**2+(YS-YT)**2)
      SINB=(YT-YS)/RANGE
      COSB=(XT-XS)/RANGE
C     B=ANGLE BETWEEN X AXIS AND GROUND LINE
      NFACET=0
C     USED TO TEST FOR DECREASING ANGLE SIZE(SHADING DETECTION)

C     BEGIN FACET FINDER INITIALIZATION
      SMALL=1.E-5
      B(1)=XS
      B(2)=YS
C     B = BEGINNING POINT
      E(1)=XT
      E(2)=YT
C     E = END POINT
      X1=B(1)
      Y1=B(2)
C     FIRST POINT IN THE SERIES
C     DETERMINE THE FACET THAT THE END POINT IS IN.
      KPT=XT+1
      LPT=YT+1
      IFLG=0

C     DETERMINE THE FACET THAT THE BEGINNING POINT IS IN.
      DO   996  I=1,2
C     DETERMINE X AND Y DIFFERENCES BETWEEN BEGINNING AND END POINTS.
      DEL(I)=E(I)-B(I)
C     IF DEL(I) GE 0, POINT(I,I)=INDEX OF NEXT FACET CLOSER TO ORIGIN
C     ALONG X (I=1) OR Y (I=2) AXIS.
      POINT(I,I)=AINT(B(I))
C     POINT(1,1)=X COORDINATE OF BEGINNING POINT
C     POINT(2,2)=Y COORDINATE OF BEGINNING POINT
      SYNE(I)=-1.
      IF (DEL(I) .LT. 0.) GOTO   996
      SYNE(I)=1
      POINT(I,I)=POINT(I,I)+1.
  996 CONTINUE

C     IF DEL(I) LT 0, POINT(I,I)=INDEX OF NEXT FACET FURTHER FROM ORIGIN
C     ALONG X (I=1) OR Y (I=2) AXIS.
      DO   994  I=1,2
C     FOR I=1 AND J=2
      J=MOD(I,2)+1
      ADEL(I)=ABS(DEL(I))
      IF (ADEL(I) .GT. SMALL) GOTO   993
      YNC (I,I)=0.
      YNC (I,J)=SYNE(J)
      GOTO   992
  993 CONTINUE
```

```
      YNC(I,I)=SYNE(I)
      YNC(I,J)=DEL(J)/DEL(I)
  992 CONTINUE
      POINT(I,J)=E(J)+YNC(I,J)*(POINT(I,I)-E(I))
      YNC(I,J)=SYNE(J)*ABS(YNC(I,J))
  994 CONTINUE
C     FINISHED FACET FINDER INITIALIZATION

  991 CONTINUE
      INDX=2
      IF (SYNE(1)*POINT(1,1) .LT. SYNE(1)*POINT(2,1)) INDX=1
C     /*WE WILL USE THE STARTMOST OF OUR TWO NEW SOLUTIONS

C     /*THIS GETS US THE ROW AND COLUMN OF THE CELL
      N=(X1+POINT(INDX,1))*.5+1.
C     N IS X COORDINATE OF FACET.
      M=(Y1+POINT(INDX,2))*.5+1.
C     M IS Y COORDINATE OF FACET.

C     IF NOT THE LAST FACET, GO TO 960.
      IF (.NOT. (M.EQ.LPT.AND.N.EQ.KPT)) GOTO   960
C        KEEP INDEXES WITHIN ARRAY
         IF(N.LT.1.OR.N.GT.40) N=1
         IF(M.LT.1.OR.M.GT.40) M=1
      POINT(INDX,1)=XT
      POINT(INDX,2)=YT
      IFLG=1
C     IFLG=1 ==> LAST FACET HAS BEEN CHECKED.

  960 CONTINUE
C        KEEP INDEXES WITHIN ARRAY
         IF(N.LT.1.OR.N.GT.40) N=1
         IF(M.LT.1.OR.M.GT.40) M=1


C     /*****************EVERYTHING TO DO WITH A FACET ONCE WE FIND IT*******
C     /**************************GOES HERE*********************************

      X2=POINT(INDX,1)
      Y2=POINT(INDX,2)
      TEMP=SQRT((X2-X1)**2+(Y2-Y1)**2)
      IF (TEMP .LT. .01) GOTO   984
      NFACET=NFACET+1
      POS(2,NFACET)=TEMP
C     /*THIS IS THE LENGTH OF THE LINE SEGMENT INTERSECTED BY THE CELL BOU
      IFACET(3,NFACET)=ITARA(4,M,N)
C     /*SURFACE ROUGHNESS
      IFACET(4,NFACET)=ITARA(5,M,N)
C     /*TERRAIN TYPE
      IFACET(5,NFACET)=1
C     /*ALL FACETS ARE INITIALLY UNSHADOWED.
      ZC=ITARA(1,M,N)*CVCM
C     ZC=AVG CENTER HEIGHT IN TSCALE UNITS
      ITILT=ITARA(2,M,N)
```

```
      JTILT=ITARA(3,M,N)
      IFACET(2,NFACET)=ITILT*COSB+JTILT*SINB
C     /*LONGITUDINAL TILT
      IFACET(1,NFACET)=-ITILT*SINB+JTILT*COSB
C     /*TRANSVERSE TILT
C     CONVERT TILT TO RADIANS
      TA1=ITILT*.001
      TA2=JTILT*.001
      XC=N-.5
      YC=M-.5
      Z1(NFACET)=ZC-(X1-XC)*TA1-(Y1-YC)*TA2
      Z2(NFACET)=ZC-(X2-XC)*TA1-(Y2-YC)*TA2
      POS(1,NFACET)=(Z1(NFACET)+Z2(NFACET))*.5
C     /*FACET HEIGHT IN TSCALES


C     **********SHADOWIING CALCULATIONS**********

C     FIRST CHECK FOR SHADOWING FROM START TOWARDS TERMINUS.

C     DO NOT CHECK FOR SHADOWIND FROM START TOWARDS TERMINUS ON CLUTTER
C     CALL TO FACET

      IF(ISHAD.NE.1) GO TO 984
      IF (NFACET .NE. 1) GOTO   985
C     DO THESE CALCULATIONS ONLY FOR THE FIRST FACET FOUND.
C     IFACET(5,1)=1
      TMIN=(ZS-Z2(NFACET))/POS(2,1)
      POS(3,1)=ATAN(2.*TMIN)
      GLINE=POS(2,1)
      GOTO   984

  985 CONTINUE
C     DO THESE CALCULATIONS FOR ALL BUT THE FIRST FACET FOUND.
      T1=(ZS-Z1(NFACET))/GLINE
      GLINE=GLINE+POS(2,NFACET)
      T2=(ZS-Z2(NFACET))/GLINE
      POS(3,NFACET)=(T1+T2)*.5

C     COMPARE ANGLES ONLY IF CHECKING FOR SHADOWING.
      IF(NOSHAD.NE.1) GO TO 984
      TMIN=AMIN1(TMIN,T1)
C     IFACET(5,NFACET)=1
      IF(T2 .GT. TMIN) IFACET(5,NFACET)=0
      TMIN=AMIN1(TMIN,T2)
C     POS(3,NFACET)=(T1+T2)*.5

  984 CONTINUE
C     JUMP TO 1 IF THE LAST FACET HAS BEEN CHECKED.
      IF(IFLG.EQ.1) GO TO 1

      X1=POINT(INDX,1)
C     /*THROW AWAY OLD X & Y
      Y1=POINT(INDX,2)
```

```
      POINT(INDX,1)=X1+YNC(INDX,1)
C     /*GET A FRESH SOLUTION (POINT(INDX,1),POINT(INDX,2))
      POINT(INDX,2)=Y1+YNC(INDX,2)
      GOTO   991

C     CHECK FOR SHADOWING FROM TERMINUS TOWARDS START POINT

1     CONTINUE
      IF (ISHAD .NE. 1) RETURN
C     DO NOT CHECK FOR SHADOWING FROM TERMINUS TOWARS START POINT ON
C     CLUTTER CALL TO FACET.


C     DO THESE CALCULATIONS FOR FIRST FACET (FACET CLOSEST TO TERMINUS).
      IFLG=1
C     IFLG=1 NOW MEANS NO FACETS VISIBLE YET FROM TERMINUS.
      GLINE=POS(2,NFACET)
      TMIN=(ZT-Z1(NFACET))/GLINE
C     AVERAGE ELEVATION ANGLE FROM TERMINUS TO LAST FACET.
      POS(4,NFACET)=ATAN(2.*TMIN)
C     COUNT BACK FROM TERMINUS TO START POINT.
      IND=NFACET-1

  976 CONTINUE
C     DO THESE CALCULATIONS FOR ALL BUT FIRST FACET.
      IF (IND .EQ. 0) RETURN
      T2=(ZT-Z2(IND))/GLINE
      GLINE=GLINE+POS(2,IND)
      T1=(ZT-Z1(IND))/GLINE
      TMIN=AMIN1(TMIN,T2)
C     CHECK TANGENTS ONLY IF CHECKING FOR SHADOWING.
      IF(NOSHAD.NE.1) GO TO 971

C     IF THE FACET WAS SHADOWED FROM THE OTHER DIRECTION, DO NOT CHECK
C     IT AGAIN.
      IF (IFACET(5,IND) .NE. 1) GOTO   974
      IF(T1 .LE. TMIN) GOTO 971
      IFACET(5,IND)=0
      GOTO 974

  971 CONTINUE
      IF(IFLG.EQ. 0) GOTO 974
      IFLG=0
C     INDICATES FIRST VISIBLE FACET FOUND.
      NFACET=IND

  974 CONTINUE
      TMIN=AMIN1(TMIN,T1)
      POS(4,IND)=(T1+T2)*.5
C     AVERAGE ELEVATION ANGLE FROM TERMINUS TO FACET.
      IND=IND-1
      GOTO   976
      END
```

```
      FUNCTION ATAN3(Y,X)
C
C     COMPUTES ARCTAN(Y/X)
C
      IF(Y.NE.0.) GO TO 10
      IF(X.EQ.0.) GO TO 20
 10   ATAN3=ATAN2(Y,X)
      IF (ATAN3.LT.0.) ATAN3=ATAN3+2*ACOS(-1.)
      RETURN
C
C   BOTH X AND Y ARE ZERO.
C
 20   ATAN3=0.
      RETURN
      END


      COMPLEX FUNCTION RICE(A,B)
      COMPLEX A
C     USING CDC RANDOM NUMBER GENERATOR.
      UR=RANF(5)*ACOS(-1.)*2.
      XR=B*SQRT(-ALOG(RANF(5)))
      RICE=CMPLX(REAL(A)+XR*COS(UR),AIMAG(A)+XR*SIN(UR))
      RETURN
      END


      FUNCTION F(T,TB)
      UF=2.3779*ABS(T)/TB
      F=1.4855-.7*UF
      IF(ABS(UF-1.).GT..001)F=COS(.5*3.1415926*UF)/(1.-UF*UF)
      RETURN
      END


      FUNCTION FD(T,TB)
      TQ=.5*TB
      T1=T+TQ
      T2=T-TQ
      FD=F(T1,TB)-F(T2,TB)
      RETURN
      END


      FUNCTION FS(T,TB)
      TQ=.5*TB
      T1=T+TQ
      T2=T-TQ
      FS=.707*(F(T1,TB)+F(T2,TB))
      RETURN
      END
```

```
      SUBROUTINE CVTXY(XFEET,YFEET,XMETER,YMETER)

**********************************************************************
*                                                                    *
*     THIS SUBROUTINE TRANSFORMS COORDINATES (IN FEET) IN A SITE-    *
*     CENTERED COORDINATE SYSTEM INTO COORDINATES (IN METERS) IN     *
*     THE REFERENCE COORDINATE SYSTEM.  IT IS ONLY NEEDED WHEN THE   *
*     TARGET'S COORDINATES ARE DEFINED USING THE SITE-CENTERED SYSTEM *
*     INSTEAD OF THE REFERENCE SYSTEM.  XEAST, YNORTH, THETA AND PHI *
*     DEFINE THE TRANSFORMATION FROM THE REFERENCE TO THE SITE-      *
*     CENTERED SYSTEM.                                               *
*                                                                    *
**********************************************************************

      COMMON/UTM/ITERSW,YNORTH,XEAST,PHI,ZMASK,XTILT,YTILT,RMS

C     THE FOLLOWING IS THE ORIGINAL UTM COMMON BLOCK.  IT INCLUDES
C     SEVERAL VARIABLES THAT ARE NEVER USED.
C     COMMON/UTM/ITERSW,YNORTH,XEAST,PHI,DRATE,IGRID(4,3),
C    $           RHMASK,XMASK,YMASK,ZMASK,IGRDR,ITERR(1717)

      XMETER=XFEET/3.28
      YMETER=YFEET/3.28

      RANGE=SQRT(XMETER**2+YMETER**2)

C   IF(OVER ORIGIN) THEN
      IF (RANGE.NE.0) GO TO 10

C   SET TO MAP VALUES
      XMETER=XEAST
      YMETER=YNORTH
      GO TO 15

C   ELSE
C       COMPUTE LOS ANGLE
 10   THETA=ATAN3(XMETER,YMETER)

C   COMPUTE X AND Y
      XMETER=XEAST+RANGE*SIN(THETA+PHI)
      YMETER=YNORTH+RANGE*COS(THETA+PHI)

 15   CONTINUE
      RETURN
      END
```

VITA

James Michael Arnett was born on June 2, 1951 in San Jose, California. He graduated from high school in Morgan Hill, California in 1969 and attended the USAF Academy, from which he received the degree of Bachelor of Science in Aeronautical Engineering in June 1973. Upon graduation he attended Undergraduate Pilot Training at Laughlin AFB, Texas, and received his wings in December 1974. He served as a C-141 pilot with the 4th Military Airlift Squadron, McChord AFB, Washington until entering the School of Engineering, Air Force Institute of Technology, in June 1978.

He is married to the former Priscilla Saulnier of Monument, Colorado. They have two daughters, Rebecca and Wendy.

Permanent address: 1046 California Street
Calimesa, CA 92320

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER AFIT/GOR/SM/79D-1	2. GOVT ACCESSION NO. AD-A083 516	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) TOWARD VALIDATION OF COMPUTER SIMULATION MODELS IN OPERATIONAL TEST AND EVALUATION		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Capt James M. Arnett		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Air Force Institute of Technology (AFIT/EN) Wright-Patterson AFB, OH 45433		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Institute of Technology (AFIT/EN) Wright-Patterson AFB, OH 45433		12. REPORT DATE December 1979 13. NUMBER OF PAGES 178
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Approved for public release IAW AFR 190-17

JOSEPH P. HIPPS, Major, USAF
Director of Information

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Computer Simulation
Validation
Experimental Design
Experimental Design in Validation of Computer Simulation Models
Radar Tracking
Missile Site Radar
Multipath Radar
Radar Clutter
Operational Test and Evaluation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Validation of complex computer simulations is considered in the context of the operational test and evaluation of the air launched cruise missile. Published literature on validation is reviewed. Validation is described as a problem-dependent process. The goal of that process is an acceptable level of confidence that the actual and simulated data agree closely enough for an inference about the simulation to be a valid inference about the actual system. The cost of accepting a given level of confidence must be balanced against the cost of obtaining a higher level. (Continued on Reverse)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

BLOCK 20.

Experimental design is seen as the key to obtaining the maximum amount of information from a limited number of test runs and is discussed. The following approach to validation is suggested. Use a screening design to identify the few, most-important variables in the simulation. Use a more complete design to specify parameter/level combinations for simulated and actual data comparison. Finally, explicitly incorporate decision analysis in judging the validity of the model based on that comparison and the use that will be made of the simulation results.

A fractional factorial design is used to screen out the important factors from the clutter and multipath sub-model in TAC ZINGER.

FILMED

5-8