CHARACTERIZATION, COMPARISON, AND VALIDATION OF ELECTROMAGNETIC MODELING SOFTWARE

E. K. MILLER
General Research Corposation
5383 Hollister Avenue
Santa Barbara, CA 93111

ABSTRACT

The continuously increasing number of electromagnetic computer models (codes) and applications thereof is one result of a rapidly expanding computing resource base of exponentially growing capability. While the growing use of computers in electromagnetics attests to the value of computer modeling for solving problems of practical interest, the proliferation of codes and results being produced increases the need for their validation with respect to both electromagnetic formulation and software implementation. But validation is perhaps the most difficult step in code development, especially for those models intended for general-purpose application where they may be used in unpredictable or inappropriate ways. A procedure or protocol for validating codes both internally, where necessary but not always sufficient checks of a valid computation can be made, and externally, where independent results are used for this purpose, is needed. Ways of comparing differing computer models with respect to their efficiency and utility to make more relevant intercode comparisons and thereby provide a basis for code selection by users having particular problems to model are also needed. These issues are discussed in this article and some proposals are presented for characterizing, comparing, and validating EM modeling codes in ways most relevant to the end user.

1. INTRODUCTION

1.1 Background

Recognition of the need to validate computer models used in various branches of science and engineering is demonstrated by the growing attention being given to this topic by various interested groups. The Acoustical Society of America, for example, has held two sessions at its annual meetings on comparing solutions for certain "benchmark" problems involving propagation of acoustic waves in the ocean, and may publish written versions of the papers in a special issue [Felsen (1988)]. Workers developing and applying eddy-current models for nondestructive testing have held a series of international workshops to compare solutions obtained for a set of specified problems [Turner (1988)]. More recently, ACES initiated a new committee on "Software Performance Standards" at its March 1988 meeting and is publishing this special issue of its Newsletter/Journal on the topic of "Software Validation." A special session on Software Validation was also organized for the 1988 IEEE AP-S/National Radio Science Meeting [Miller (1988)] and AP-S AdCom approved forming a new committee on the general topic of "Electromagnetic Modeling Software." The presentations made at that session are listed under references.

This article discusses my own perspective on this very important topic and provides a "zeroth-order" starting point for developing a systematic procedure or *protocol* for the general problem of characterizing, comparing, and validating EM modeling software. It is hoped that this and other articles in this special issue will stimulate further discussion and lead to specific recommendations and procedures for such a *protocol.*

1.2 Essential Attributes of a Computer Model

Among the various attributes that might be selected for characterizing and comparing computer models for electromagnetic or other applications, the following are most fundamental:

Accuracy/Reliability--Above all else, a modeling computation must posses acceptable, preferably known, and better yet "dialable" accuracy. This is an attribute to which all others, however desirable they might be, must be considered secondary, for invalid results have no value and can even be detrimental.

Efficiency/Productivity--Following accuracy as a desirable attribute is efficiency, both with respect to the computer resources needed for the modeling computation as well as with respect to the human resources needed to develop the input required to run the model and to interpret the results that the model produces. While it is obvious that a model that efficiently produces inaccurate results has no value, it is equally true that a model that produces acceptably accurate results but which requires computer and/or human resources incommensurate with the application has little more value.

Utility/Applicability--Finally, we must consider a model's utility in terms of the kinds of problems to which it can be applied. At this stage, general-purpose modeling codes and specialized, single-problem codes become most differentiated. On the one hand, it is always easier to develop a model specialized for a particular problem that will be more accurate and/or efficient than a general-purpose code that can model that same problem. On the other hand, the more widely applicable a given modeling code becomes, and the more easily used it is, the greater its utility for the non-specialist who does only infrequent modeling.

The purpose of the discussion is to suggest development of an experimental validation *protocol* for characterizing electromagnetic modeling software and assessing its performance with respect to these attributes. The word "experimental" is used here to emphasize the need of literally performing computer experiments to obtain the needed data, since the ultimate test of modeling software can only be done using the numerical results that it produces. In section 2 we first discuss the basic, universal steps involved in developing a computer model and point out the need for systematic and accepted procedures of quantitatively determining model validity, a need made all the more critical by the growing amount of modeling that is being done. Discussed in section 3 are some of the factors that need to be considered in model validation, and in section 4 we elaborate on the roles of internal and external checks used for this purpose. This idea is developed further in section 5 in terms of suggesting primary and secondary "benchmarks" for validation, in a fashion analogous to how the National Bureau of Standards does so for metrological needs. We choose the word "benchmarks" rather than "standards" as used by NBS because the latter implies a formal, legalistic definition within such organizations as the IEEE that does not seem appropriate at this point in connection with EM modeling software. Error types and error sources are examined in section 6. In section 7 we consider in more detail the components of model characterization and comparison in terms of the above attributes, and in section 8 we close with some concluding remarks.

2. DEVELOPING AN EM COMPUTER MODEL

The process of developing a computer model involves a small number of basic steps whatever particular details are involved. Because these steps are universally encountered, it is worthwhile to summarize them as follows:

Conceptualization--Encapsulating observation and analysis in terms of elementary physical principles and their mathematical description (e.g., the idea that a Green's function can be used to represent the fields of arbitrary source distributions).

Formulation--"Fleshing out" of the elementary description into a more complete, formally solved, mathematical representation (e.g., development of an integral equation from a source integral and required boundary conditions on field behavior).

Implementation--Transforming the formulation into a computer algorithm using various numerical techniques (e.g., using the method of moments).

Computation-- Defining the model and obtaining quantitative results (the "crank-turning" stage).

Approximation--Simplifying operations and assumptions that can arise at any step in the process of model development and application (e.g., reducing a surface integral to a line integral by using the "thin-wire" approximation).

Validation--Determining the numerical and physical credibility of the computed results (using analytical, experimental, and/or other numerical results for comparison with the computer model being employed).

For any applications-oriented software, the most time-consuming and laborsome of these steps is the last, that of validation. This is becoming increasingly the case as growing computer power has expanded the size, complexity, and volume of problems that are routinely modeled. Whereas computer resources available in the 1960s limited the amount of data needed to describe problems and represent the results, there has been an explosive growth in the scope of applications as summarized in Table I.

Long after work on the model has been completed, questions will continue to arise about its performance. Such questions include: is a given result valid; can the model be used reliably for a given problem; what might be the numerical accuracy and physical relevancy of the results that are obtained? These questions become especially important as modeling moves from a primarily research environment to one which involves an increasing emphasis on analysis and design applications. The difficulties caused by uncertainty over model performance also increase with the expanding proliferation of modeling codes and computational resources becoming available. The question of perceived model validity is of particular concern in that it can lead to correct results being rejected because of unwarranted skepticism or acceptance of incorrect results because of misplaced confidence. Either outcome is undesirable and both should be avoided by developing the validation procedures needed for an appropriate level of user confidence to be achieved.

Without essentially "exact" results to serve as benchmarks, there will always be some lingering doubts regarding the validity, let alone accuracy, of computer models. Unfortunately, as is well known, there are few closed-form, exact solutions available from classical electromagnetics. For a 3D computer model to match results for a spherical body is hardly convincing anymore that the same model will work as well for a more arbitrary body geometry. Checks of the kind provided by the sphere can be regarded only as necessary, but not sufficient, conditions for solution validity. But without reference solutions to provide benchmark results, quantification of computer-model accuracy and validity for the most part will continue to remain an open question.

3. AN APPROACH TO VALIDATION

It might be helpful at this point to clarify what validation should mean in a modeling context. Among the definitions given by the 1975 American Heritage Dictionary are the following:

--to declare or make legally valid;

--to mark with an indication of official sanction;

--to *substantiate* or *verify*,

with the latter being most relevant to our discussion. Proceeding with the understanding that validation means to substantiate or verify model performance, several key issues arise as are discussed briefly below.

3.1 What Problems and Solutions?

The first decision to be made in model validation is the selection of appropriate test problems. Although numerical validation can (and should) be performed within a model itself using internal checks as discussed below, a more logical starting point for model validation is the use of external data or checks. From an analytical viewpoint, we need to consider for which problems are answers, preferably of known accuracy, available to serve as independent sources of results? Alternatively, are there certain kinds of problems for which experimental data of needed accuracy can be obtained? Furthermore, from among that set of candidate analytical and/or experimental problems, which provide the most appropriate testing of model capabilities? For example, while one way of validating a wire code such as *NEC* can be to compare results for scattering from a wire-mesh model of a sphere with the MIE series, that would not be especially relevant for determining the code's performance when used for modeling wire antennas. As discussed further below, we suggest that the problems and solutions that are selected for validation purposes might be usefully assigned to one of two categories, described as primary and secondary benchmarks.

3.2 What Comparisons?

A second decision concerning model validation involves which quantities are to be used for this purpose. Most obviously, these quantities should include physical observables for which measured data, at least in principle, can be obtained. But mathematical quantities and relationships which might be essentially inaccessible using measurements might also be useful as sources of data for validation purposes. For example, the eigen values and eigen vectors of the moment-method impedance matrix might be candidates for use in model validation, but they are not directly measurable in general.

3.3 How Accomplished?

Finally, we must carefully consider how the quantities chosen for use in model validation are to be compared, and over what range of variables and parameters the comparisons should be performed? The spatial or temporal variation of a given field quantity might be appropriate for some applications. On the other hand, a result derived from integrated measures of such quantities, for example total scattered power, might be more relevant for other purposes. The former approach might be called microscopic in that the fine structure of the solution is being examined, while the latter could be described as macroscopic because it provides a less detailed but broader means of comparison.

These various issues are not easily settled and will require thoughtful consideration and, most likely, systematic refinement as procedures for model validation evolve. We are suggesting essentially that an experimental *protocol* be developed for this purpose. This *protocol* would set down clearly defined procedures for validating present and future models in an agreed-upon way that is both physically and numerically relevant to intended applications.

4. INTERNAL AND EXTERNAL VALIDATION CHECKS

We noted above that essentially two kinds of procedures can be used to establish some quantitative measure of code validity:

1) **Internal Validation**, a check that can be made concerning solution validity within the model itself; and

2) **External Validation**, a check that utilizes information from other sources which could be analytical, experimental or numerical.

4.1 Internal Checks

Existing computer models often do not perform internal checks on the results they produce, but instead leave that as an exercise for the user. For example, *NEC* could provide and indeed has been exercised to give various kinds of checks relating to pow balance, reciprocity and boundary-condition matching. But the software to perform the wide range of internal checks that might be most useful for a given application is not completely implemented in the code, but instead may need to be "patched in" by the user for a particular problem and check. It would be extremely valuable if a variety of such checks could be built into the code and exercised as desired by the modeler. Some possibilities for internal checks are outlined in Table 2.

As a particular example of the possible applications of internal checks, consider the case when a problem new to the modeler is encountered and the initial results are obtained. Present practice usually involves "eye-balling" the data to see if it "feels" right, perhaps having first run some documented test cases to verify code performance. Since these test cases would not likely resemble the new problem, their successful solution might not provide much insight concerning the new results. If, however, a series of checks built into the code could then be exercised at the modeler's discretion to verify that conditions necessary for a valid solution of Maxwell's Equations are satisfied, confidence in the model's numerical reliability could be more readily established. These checks might range from fairly exhaustive, such as computing boundary fields to determine how well boundary conditions are satisfied, to fairly simple, such as evaluating the degree to which reciprocity and power conservation are demonstrated. They could only be viewed as necessary but not sufficient conditions for solution validity, and could only involve such behavioral aspects as are not implicit in the model already (e.g., some formulations produce symmetric matrices so that bistatic scattering and transmit-receive reciprocity are assured analytically). Developing a figure-of-merit from the results of such checks that would provide a "quality factor" (or more if application-specific measures are useful) for the solution in a single number seems not only feasible but highly desirable.

4.2 External Checks

The second kind of check involves use of independent data from other sources. Perhaps the most convincing overall is experimental data, but analytical or numerical results should be comparably useful. Indeed, one of the most convenient computational checks would be provided by a code that permits two different numerical models to be developed for the same problem, for example by incorporating user-selectable basis and weight functions. For greatest utility, such checks ideally should not be microscopic or of single-point nature, e.g. a comparison of results for input impedance at a single frequency. This is because experience shows that computer models produce results that exhibit apparently slight frequency shifts, angle shifts or spatial shifts in field quantities with respect to "exact" solutions, or even other computer models. When the effects of such shifts are observed near maxima in the response of interest, they may appear relatively insignificant, but when examined near deep minima or even nulls, the differences can become unbounded. Consequently, macroscopic or global comparisons are usually more meaningful, but even they may not be straightforward to interpret. If the shifts mentioned are observed, it would seem more appropriate to develop a correlation measure such as computing the minimum squared difference between the two results as they are shifted along the axis of the common variable, rather than simply doing an absolute differencing. For other models and applications, the results may be even less directly comparable, as is the case for IE -and DE-based models. Some work is needed in the general area of how results from two different representations of the same problem can be most meaningfully compared. In Table 3 are summarized some of the external checks used for model validation.

5. DEVELOPMENT OF VALIDATION BENCHMARKS

A model for development of a validation *protocol*, using the term introduced above, might be provided by how the National Bureau of Standards (NBS) has established standards for various metrological applications. NBS undertakes to develop "primary" standards as a means of providing the most accurate metrological benchmarks for various fundamental quantities. They also develop a set of "secondary" standards whose accuracy is derived from the former but which have the advantage of being more portable and less expensive to employ. If a set of standards or benchmarks were to be developed for checking computer models using a prescribed methodology along similar lines, more confidence could eventually be placed in a model that satisfied certain testing criteria, i.e., the *protocol*.

The idea of using benchmarks in the computer world parallels what has been done for years in the experimental one. Radar scattering ranges, for example, have been routinely calibrated using a metal sphere as a reference target. Even if the RCS of the sphere were unknown, the measured results could have been calibrated with respect to this basic target. Since the sphere was the first 3D target whose cross section could be quantified in absolute terms, it enabled absolute results to be inferred for targets being measured. What has worked so well in the experimental world is worth examining to asses how it might contribute to the calibration and validation of computed results.

Use of the word "calibration" might appear inappropriate in a computational setting. It should be noted, however, that computing the effects of parameter variations for a given model is likely to indicate more accurately relative differences among the various models than any of the individual models are correct in absolute terms. In such situations, experimental data might actually be helpful in "calibrating" a numerical model.

5.1 Primary Benchmarks

For primary benchmarks, we might expect to use those problems for which solutions of essentially arbitrary numerical accuracy are available. These could include classical, separation-of-variables solutions for conducting and penetrable objects such as the circular and elliptical infinite cylinders in two dimensions, and sphere, spheroid and ellipsoid in three dimensions. Other, basically analytically solvable problems, even though requiring substantial computation to obtain numerical results, could also serve as primary standards. The slotted infinite circular cylinder and the spherical shell with a circular aperture, are examples of such problems.

5.2 Secondary Benchmarks

By definition, the analytical solutions which would serve as primary benchmarks are limited to a set of special problems. Of necessity, the next level of validation would rely on derived or secondary benchmarks by using computer models whose validity has been quantitatively established for a set of carefully defined problems. The codes used as secondary benchmarks might be generally applicable to a large variety of problem types but would be proposed as benchmarks only for those problems for which their quantitative accuracies had been specifically established.

6. KINDS OF MODELING ERRORS

Discussion thus far has focussed on modeling errors without considering the kinds of errors that can occur and what can cause them. We define error types and error sources below.

6.1 Types of Errors

The error type refers to the general effect produced on the modeling process as:

Type 0 errors--Type-0 errors keep a program from running to conclusion, and are therefore the most obvious when they occur, but not necessarily the easiest to correct.

Type 1 errors--Type-1 errors occur when a program runs to conclusion to produce the requested output which contains obviously incorrect results. A fairly common example is that of obtaining a negative input resistance for an antenna.

Type 2 errors--Type-2 errors arise when the program runs and produces what appear to be physically plausible results, but which are invalid for the problem being modeled. This category of error is generally most insidious because it is generally the most difficult to identify and correct.

Type 3 errors--Type-3 errors are user dependent as they occur when the modeler misinterprets, mistrusts, or misuses the results produced in the computation. It is reasonably well accepted for example, that computer models produce results that are generally more accurate on a relative than on an absolute basis. For example, often the nulls and peaks of a radiation pattern or a transfer function are shifted between computation and measurement, although their overall structures may be essentially the same. A modeler unacquainted with such shifts might consequently not accept computed results which exhibit them, even though they are basically correct and useful for the problem under consideration.

6.2 Sources of Errors

Modeling errors can be caused in at least four ways:

Software errors--These can originate from the operating-system software or programming errors in the modeling code.

Numerical modeling errors--A numerical modeling error arises from obtaining insufficiently accurate numerical results for the model that has been selected, one example being non-converged results. Another example is that of using word sizes of insufficient bit length which affect matrix-fill and matrix-solution accuracy due either to machine limitations or user preference.

Physical modeling errors--A physical modeling error arises from an inadequate "match" between the physical reality of interest and the numerical model that has been used. A common example in antenna modeling is that of improperly representing the source region in the numerical model, giving rise to an error which primarily affects the input susceptance.

User errors--Aside from such obvious sources as input-data errors, this category includes misapplication of the model by violating stated limitations intrinsic to the formulation or its numerical implementation. An example of the former is use of a model based on the magnetic-field integral equation for open or thin structures. Violation of the thin-wire approximation by using segment lengths shorter than the wire diameter is one example of the latter. User errors can occur because the model results are misinterpreted, mistrusted or misused due to unrealistic expectations, unwarranted skepticism or blind faith. The two extremes of user reaction that can follow are rejection of correct results or acceptance of wrong results, either of which might be equally unfortunate.

7. MODEL CHARACTERIZATION AND COMPARISON

From the user's viewpoint, the major differences which affect selecting from among alternate available modeling codes are their suitability for the problem of interest, the costs (computer and human) of obtaining the desired solution using a given model, and the accuracy that the model provides relative to the problem requirements. Many of the factors which preoccupy code developers are of little direct concern to users and are preferably transparent in application except as these factors affect accuracy, efficiency, and utility. These developer issues nevertheless can be of crucial importance when attempting to validate and otherwise compare different models for a common set of test problems. Because so many analytical and numerical choices must be made in the course of model development with respect to the solution domain (e.g., time or frequency), the field propagator (e.g., the Maxwell curl equations, a Green's function, modes, or ray optics), and the numerical treatment (basis and testing functions and the solution procedure), model intercomparability is not always straightforward. Meaningful ways of relating results from two or more models are therefore required beyond superficial comparisons of, for example, antenna input impedance or scattering cross section versus frequency. These kinds of comparisons provide useful but incomplete model intercomparison.

In the formative days of electromagnetic computer modeling, much attention was devoted to the formulations and numerical methods employed in developing an algorithm. Lengthy debate took place in the numerical sessions held at various symposia regarding what integral equation was the "best" and what expansion and testing functions gave the fastest convergence, the most accurate results, the minimum computer time, etc. As more experience was gained in using computer models, it had to be admitted that each approach had its own special advantages in some limited region of a multidimensional parameter space, and that generally there was no one model that was best everywhere. It became increasingly evident that the basic requirements for a particular model to have value were that it must be mathematically and numerically valid, and that its strengths and limitations with respect to its formulation and numerical treatment be understood and taken into account in application. The result was that more attention was paid to general applications than to the specific approach and to exploring the practical bounds of what could be expected from computational electromagnetics.

During the past few years, some aspects of that earlier experience have begun to recur. There have been energetic discussions regarding the ways by which matrices can be best solved on the numerical side (e.g., LU decomposition, conjugate-gradient iteration), and in the formulation area about whether integral-equation (IE) or differential-equation (DE) models are better and whether DE models based on finite-difference or finite-element approaches should be preferred. Such questions are important with respect to advancing the computational State-of-Art (SOA), the outcome of which will greatly affect the realm of what is practical in applications. It would be extremely beneficial if periodic, informed assessments of the SOA which address such questions could be made in an objective fashion, understandable to a non-specialist user of the model. All of these questions relate to the general problem, aside from solution accuracy, of how different models can be most meaningfully compared or described with respect to accuracy, efficiency, and utility.

Such assessments could be organized with respect to these basic attributes. Although accuracy or validation is of principal concern here, a complete model characterization requires that all three be considered. We discuss each in turn below.

7.1 Accuracy

Application-relevant, model-independent measures must be developed for quantitatively assessing the accuracy of computed results. As a starting point, we might consider the following categories of results:

1) Far-field quantities--For exterior problems, far-field quantities and results derived therefrom are often the primary goal of the model application. These include macroscopic or integral measures such as total far-field power as well as the microscopic or angle-dependent results from which these quantities are derived. In those cases where the far-field polarization properties are important, these quantities are needed separately for the appropriate field components. Because the E and H fields are related simply by the medium wave impedance, only one of them must be dealt with explicitly.

2) Near-field quantities--The near fields are generally thought to provide a more demanding measure of model performance than does the far field. Because the E and H fields are, in the near field, related by a position-dependent impedance, both are relevant quantities for validation purposes. We note that for interior problems, all fields, by definition, are near-field quantities in some sense.

3) Boundary quantities--These are quantities associated with steps in medium properties at surfaces on which boundary conditions are stated as part of the problem definition. Both the fields themselves, as well as their associated sources, are boundary quantities that are useful for validation. Derived quantities such as antenna input admittance also fall into this category.

4) Other quantities--There are other numerical quantities provided by many models for which no direct physical measurement can be made but which nonetheless are useful for assessing computation accuracy. For example, convergence of the eigenvalue (EV) spectrum as the number of unknowns is increased in an IE-based model serves as one measure of convergence and by inference, of accuracy. But the EV would not be directly measurable, although it might be obtained from computation based on the appropriate physical measurement.

5) Approximations--Affecting the accuracy of all the above measures are the approximations inherent in any model whether made in the formulation or subsequent numerical implementation. These approximations affect both the kinds of problems for which the code can be used and the level of accuracy that might reasonably be expected from it, and can also dictate how the code might be used for achieving increased accuracy for problems that stretch its capabilities. For example, a model that can be applied to geometries having edges or bends but which provides no special treatment for such features might yield better results when the sampling density is systematically increased in such regions.

7.2 Efficiency

As mentioned above, from the viewpoint of the overall cost of achieving a specified accuracy for the electromagnetic observables needed from a given model, efficiency involves two components: computer and human. The computer cost itself might be stated in several ways including total CPU time or the money charges of the computation. A more relevant and hardware-independent measure of the computing cost would instead be provided by estimating the total number of floating-point operations (FLOPS) required for the overall computation. It might be even more informative to multiply this number in turn by the number of bits manipulated per FLOP to establish the total number of bit operations, or BLOPS, required for the model computation to be accomplished to some acceptable accuracy.

The BLOP count could be especially significant in comparing two otherwise similar models applied to the same problem when one requires higher-precision computations because it is less well conditioned. The FLOP count, on the other hand, could be more relevant when one of the models produces unused information relative to the other. An example of the latter situation is that of computing the factored form of the impedance matrix when using an IE model for an antenna problem where a solution is needed for only one excitation, in which case an iterative procedure might be more efficient.

To summarize, among the factors which affect efficiency are:

1) Storage efficiency--These include both program (S_P) and variable storage (S_V). Generally speaking, as problem size and complexity increase, variable storage will eventually predominate.

2) Model efficiency--Model efficiency is determined by a number of components as outlined below.

a) Variable and equation sampling--A choice that affects both model accuracy and efficiency is that of the basis functions and testing functions used to approximate the physical quantities being sampled. Subsequent computation of physical observables can also be sensitive to the basis and testing functions.

b) Matrix fill and solution times--While using more complex basis and testing functions can increase the fill time for an IE system matrix, this is at most an N^2 process, the cost of which will always be exceeded by the solution time as N increases. Depending on the numerical condition number, computing the matrix coefficients might require greater accuracy, or the matrix solution will need to be done using greater precision.

c) FLOP and BLOP count--The total number of multiply/divide operations and their associated bit lengths which are necessary to model a given problem to acceptable accuracy provide a fundamental measure of model performance. While it is accepted that the computer-time dependence on the problem size in wavelengths or total number of unknowns is useful to compare different models, the required precision of the computations also needs to be considered.

3) Information efficiency-- Although all electromagnetic models proceed from Maxwell's equations, their particular formulation or numerical implementation may result in substantial differences in the kind and amount of information provided by the basic computation. A DE-based model, for example, provides the spatial fields throughout some solution volume whereas an IE-based model normally yields only the sources over bounding surfaces. Matrix solution by factorization provides a right-hand-side (RHS) independent solution matrix whereas iteration requires that an entirely new solution be computed for each new RHS. It might be useful to define a measure of *information efficiency* for given models applied to given problems. One possibility would be that of dividing the information concerning electromagnetic

observables actually needed in the application by the total information given by routine use of the model. For antenna problems which involve a single point of excitation but which are modeled using an IE-based code which solves the impedance matrix by factorization, the efficiency would be of order 1/N.

4) User efficiency--This aspect of efficiency is not so dependent on the formulation, numerical implementation, and solution procedures as it is on the user interface provided by a model's developers. Thus, a model which might be deficient with respect to user efficiency but which is otherwise attractive could be improved with the addition of more "user-friendly" interfaces. For a user choosing among competing models and having less demanding applications, user efficiency could be the most important overall factor in making a selection.

5) Overall efficiency--Each of the above measures of efficiency might be combined in an appropriately weighted sum to derive some overall efficiency measure, the weights being determined by the importance of each component to a prospective user of a given code.

7.3 Utility

This attribute is perhaps less easily defined than accuracy or efficiency but can be regarded as including all those factors not included in either. Utility essentially measures the kinds of problems for which a model might reasonably be used. Among the factors which comprise utility are:

1) Geometrical configuration--Problem geometry can be described in terms of the dimensionality of the solution space (1D, 2D, 3D), the permitted geometrical characteristics of the problem (straight wires, strips, flat plates, curvilinear shells, smooth closed bodies, bodies having discontinuous normal derivatives, etc.), interior or exterior, etc.

2) Electrical characteristics--These embody the object's description in terms of its electromagnetic parameters among which are size in wavelengths, surface impedance if primarily impenetrable or spatial permittivity and permeability if penetrable and possibly inhomogeneous, whether nonlinear or time varying, etc.

3) Excitations--Most modeling involves excitation either by an incident plane wave if a scattering application, or a localized voltage (or current) source if an antenna problem. Other excitations could also include Gaussian beams, nearby current sources, diffracted fields described by GTD, or fields given by modal expansions.

4) Solution domain--IE- and DE-based models are available for both the frequency domain and the time domain. Although they have a substantial intersection of applicability, there are classes of problems for which one provides the more natural and efficient approach. Problems which involve time-varying and nonlinear media or objects are generally more efficiently handled by a time-domain model. On the other hand, problems which include frequency-dependent effects such as those caused by a conductive medium like a plasma, are generally more efficiently modeled in the frequency domain.

5) Input/Output requirements--All modeling codes share the common requirement that the "real-world" problem of interest be expressed in "model-world" terms

compatible with that particular code's formulation and numerical implementation. But the user assistance provided within the code to transform from the real world to the model world varies widely from code to code, as does the presentation of results upon completion of the computation. These differences can be extremely important to prospective users when selecting from among the various alternatives that might be available.

6) Hardware requirements--While code transportability seems to be improving due to the increasing use being made of structured programming, standardized languages, and standardized subroutine libraries, it is still important for a prospective user to know which machines and configurations have been used to run a given code. Especially important in this regard is knowledge of such machine-dependent issues as the minimum word size or the computation precision required to obtain acceptable accuracy for the various kinds of problems for which the code is intended.

8. CONCLUDING REMARKS

In this discussion we have examined a number of issues involved in characterizing, comparing, and validating EM computer models. We pointed out that the three principal attributes of accuracy, efficiency, and utility determine the overall value of a computer model to a prospective user. We then suggested that an experimental validation *protocol* is needed for characterizing electromagnetic modeling software and assessing its performance with respect to these attributes. To establish a point of departure for developing such a *protocol*, we first discussed the basic steps involved in developing a computer model such as conceptualization, formulation, approximation, implementation, computation, and validation. The need for systematic and acceptable procedures of quantitatively determining model validity was emphasized, a need observed to be made all the more critical by the growing amount of modeling software that is being developed and modeling that is being done. We next discussed some of the factors that must be considered in achieving model validation in terms of what problems and solutions might be selected, what comparisons should be made, and how this might be accomplished. We contrasted the roles of microscopic or detailed, and macroscopic or integrated, error measures, elaborating in particular on the use of checks that might be made internal or within a modeling code, and external checks which employ independent analytical, numerical, or experimental data. The idea was developed further in terms of suggesting primary and secondary "benchmarks" for validation in a fashion analogous to how the National Bureau of Standards does so for metrological purposes. Error types and error sources were next examined, and the components of model characterization and comparison in terms of the above attributes were then considered in more detail. The basic point we wish to stress is that computer modeling in electromagnetics has matured to a stage where an appropriate proportion of attention must be devoted to assessing model performance in terms of application-relevant measures within the overall context of accuracy, efficiency, and utility.

9. REFERENCES

(All those listed below were presented June 9 during the 1988 IEEE AP-S Symposium at Syracuse University at the Special Session on Software Validation).

L. B. FELSEN, **Benchmarks: An Option for Quality Assessment,** POLYTECHNIC UNIVERSITY

S. J. KUBINA, C. W. TRUEMAN, **Some Practical Considerations in the Validation of EM Modelling Codes,** CONCORDIA UNIVERSITY

A. C. LUDWIG, **Validating Scattering Calculations Using Boundary Value and Internal Field Checks,** GENERAL RESEARCH CORPORATION

E. K. MILLER, **Development of EM Modeling Software Performance Standards,** ROCKWELL SCIENCE CENTER

Y. RAHMAT-SAMII, J. MUMFORD, R. THOMAS, **Reflector Antenna Distortion Compensation by Array Feeds,** CALIFORNIA INSTITUTE OF TECHNOLOGY

H. A. SABBAGH, L. D. SABBAGH, J. R. BOWLER, **A Model of Ferrite-Core Probes Over Composite Workpieces,** SABBAGH ASSOCIATES

T. K. SARKAR, E. ARVAS, **What Do You Mean by a Solution to an Operator Equation?,** SYRACUSE UNIVERSITY

L. TURNER, **Workshops and Benchmark Problems to, Validate Eddy Current Computer Codes,** ARGONNE NATIONAL LABORATORY

A. D. YAGHJIAN, M. G. COTE, M. B. WOODWORTH, **Progress Toward a Benchmark Solution for 3D, Perfectly Conducting Scatterers with Edges and Corners,** RADC, HANSCOM AFB

Table 1

GROWTH IN EM MODEL CAPABILITIES AND CHARACTERISTICS FROM THE 1960s TO THE PRESENT

	~1960	NOW
NUMBER OF UNKNOWNS	A FEW 10's	1,000's to 1,000,000's
PROBLEM COMPLEXITY	CANONICAL SHAPES, SIMPLE DEFORMATIONS	SHIPS, AIRCRAFT, ETC.
MATERIAL COMPOSITION	PERFECT CONDUCTORS	LOSSY, ANISOTROPIC, INHOMOGENEOUS
FORMULATION	INTEGRAL EQUATIONS	INTEGRAL EQUATIONS, PARTIAL DIFFERENTIAL EQUATIONS, HYBRID TECHNIQUES, ETC.
SOLUTION DOMAIN	FREQUENCY	VARIOUS

Table 2
INTERNAL CHECKS USEFUL AS MEASURES OF SOLUTION VALIDITY

CONVERGENCE MEASURES

MEASURE	**EXAMPLE**	**TESTS**	**REMARKS**
Local	$\lim[I(s)]$, $\lim[\mathbf{E}(\mathbf{r})]$ as N-->N_{max}	Convergence of input impedance, current, fields, etc.	Reasonable measure of solution behavior, but can yield non-monotonic result.
Global	$\int I(s)I^*(s)ds$ or $\int \mathbf{E}(\mathbf{r})\mathbf{E}^*(\mathbf{r})dx^n$	Convergence over entire object of current or convergence of field in n=1,2, or 3 dimensions.	A more complete, but less sensitive measure of convergence.
Random (local or global)	$\sum F(\mathbf{r}_n)$, with $F(\mathbf{r}_n)$ a field quantity which is function of a random variable $\mathbf{r}_n$.	Convergence of any field quantity measured by a random observation variable.	Permits estimation of convergence and uncertainty of convergence estimates.
		OTHER MEASURES	
Power Balance	$P_{in} + P_{loss} = P_{radiated}$	Whether supplied power equals sum of radiated plus dissipated power.	Provides good check on antenna source model for radiation resistance. A necessary, but not sufficient condition.
Boundary Condition Matching	$E_{tan}(\mathbf{r}') = 0$, $\mathbf{r}'$ on object modeled	Degree to which specified conditions on the boundary are satisfied.	Most fundamental check on solution. Consistency requires use of same weight function as for model itself (to obtain whatever smoothing the weight function provides). Can be computationally expensive. Necessary and sufficient condition.
Reciprocity	$E(\varphi_1^{inc},\varphi_2^{scat}) = E(\varphi_2^{inc},\varphi_1^{scat})$	Whether interchanging observation and source locations yields identical results.	Useful check for antenna and bistatic scattering patterns. Necessary but not sufficient condition.
"Non-physical Behavior" of Solution	—	Whether computed results exhibit physically reasonable behavior.	Can be a subjective check. One example is provided by spatial oscillation in current when thin-wire approximation is violated.

Table 3

EXTERNAL CHECKS USEFUL AS MEASURES OF SOLUTION VALIDITY

MEASURE	EXAMPLE	TESTS	REMARKS
		ANALYTICAL	
Any observable provided by a formally exact solution.	Sources, near and far fields.	Any observable provided by the computer model.	Provides a necessary and sufficient condition for solution validity. Available for only special geometries, but gives canonical benchmarks.
		COMPUTATIONAL	
Far fields.	Radiation pattern, bistatic and monostatic scattering pattern.	Consistency of the quantity least sensitive to solution errors.	A useful test, but one which is often subject of angle shifts between results from two models.
Near fields and sources.	Near-field cuts, current and charge distributions.	Quantities most often directly computed by model.	A more demanding test for comparison, but one which often exhibits spatial shifts between models.
Input impedance/ susceptance.	---	Source models and single-port input characteristics.	Especially sensitive measure in terms of input susceptance. Highly advisable to examine over a range of frequencies because shifts in frequency also occur.
		EXPERIMENTAL	
Same observables as used for computational checks.	---	Physical modeling error and relative correlation of actual problem with numerical model.	Perhaps the most reassuring check to make, but also often the most difficult.