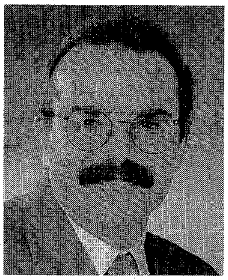

John L. Volakis
Rad. Lab., EECS Dept.
University of Michigan
Ann Arbor, MI 48109-2122
(313) 764-0500
(313) 747-2106 (Fax)
volakis@umich.edu (email)

Forward by JLV

Engineers rely, to an increasing degree, on computational results, for developing designs and for understanding the response of electromagnetic systems. If the computational algorithms are to be accepted by the industry, it is necessary that software developers provide the user with the means to evaluate the error bounds of the results delivered by the computational models. Apparently, the computational scientists and engineers in mechanical and applied mechanics, who are often concerned with the design of critical components, have been incorporating error estimation and criteria into their design for some time. In electromagnetics, the issue of error prediction and control has only been recently discussed. One of the reasons for the recent interest in error analysis and control is due to demands from the user community. General-purpose software packages for electromagnetic analysis are now routinely used by practicing engineers, and by students in the classroom. Clearly, any computational result must be accepted on the assumption of some known error bounds, if it is to be applied in practice. However, we rarely (if ever) see that such error bounds are provided along with the answer by any of the available EM software. A second reason for a need to generate error bounds is prompted by our aims to use software for designing complex and large systems that would otherwise be prohibitively expensive to build and test in the laboratory. That is, EM software tools must be capable of the analysis and design of final products in a reliable manner. Clearly, this cannot be accomplished without the availability of error bounds.

Error analysis and control are, of course, inherent features of every successful numerical method. A posteriori estimation by a software package can be done, but controlling it is not easy! The latter requires further research and development, which is expensive, time consuming, and may not likely increase the selling price of the software. Nevertheless, error analysis and control for large-scale simulations is necessary, and the scientific community must address it effectively over the next few years.

We begin this discussion in the EM Programmers Notebook by introducing the views of two authors on error analysis. One of the contributors (J. Rautio) presents the issue from a user's point of view, whereas the other contributor (D. Estep) is a mathematician. In the latter paper, Dr. Estep and his co-authors give us an appreciation of the difficulty associated with error estimation in finite-difference time-domain solutions. I want to thank both of them for providing these contributions, and for putting up with my "occasional" criticisms of their papers.

By the time this issue of the *Magazine* appears in print, the Applied Computational Electromagnetics Conference (to be held in Monterey, CA, USA) will be over, and the *Proceedings* of this conference will include several papers on error analysis, estimation and control. These papers will contain a mix of viewpoints from mathematical and engineering points of view, and I encourage you to review them. Meanwhile, we will be very interested in receiving contributions on the subject for inclusion in the EM Programmers Notebook.

The Microwave Point of View on Software Validation

James C. Rautio
Sonnet Software, Inc.
135 Old Cove Road, Ste. 203
Liverpool, NY 13090
Tel: (315) 453-3096

1. Introduction

Validation [1] generally consists of a "Good Agreement Between Measured And Calculated" (GABMAC) plot. Unfortunately, when one looks at microwaves-related papers, error bars are rarely plotted on measured data, and quantitative estimates of error are typically absent. Further, convergence analyses, if performed, are rarely mentioned, and test structures are not checked for sensitivity to analysis error. Meanwhile, the discussion of these issues in the antenna field is on-going and active. In fact, it has progressed to the point that you are asking for input from other technical areas. I will do my best to describe the situation in the microwave field as I see it.

There is now considerable commercial microwave-electromagnetic software in use in applied situations. Thus, the comments I make apply to both researchers and to vendors. The primary difference between the two, with regard to software validation, is that one seeks the approval of their customers, while the other seeks the approval of reviewers and colleagues. Both situations can be described in terms of "sales," in which validation plays an important role.

First let's discuss GABMAC. All papers (or sales flyers) I see invariably report excellent agreement. However, the authors can not possibly know if the agreement is excellent. Why not? Because in order to decide what is "excellent," you must have a specification. The only person with a specification is the design engineer, working on a specific project. In some cases, ±10% is excellent; in other cases, ±0.1% isn't good enough. With the definition of "excellent" varying over three orders of magnitude, how can the author of a technique possibly know what is excellent? Simple: they define "excellent" to be whatever their technique happens to provide.

2. Error definition

Rather than letting the researcher or vendor define excellent, the designer needs to be provided with a quantitative value for the error. Then, the design engineer can judge if the agreement is "excellent," with respect to his needs.

The problem is, what should that number be, and how do we get it?

Most microwave engineers use scattering parameters (S parameters), normalized to 50 ohms. So, the obvious choice is to specify percent error with respect to S parameters. As an example, let's calculate the error for the reflection from a perfect, 50-ohm load. The percent error is 100 × (correct - calculated)/correct. The correct value for the reflection coefficient is 0. We now have a small problem. No matter how much we try, we can not use S parameters to calculate percent error. Smith-chart error circles also prove inconvenient, because they can extend outside the Smith chart. Most passive analyses never provide such results, even when the error is large.

We have found it useful to define percent error in terms of the values of the lumped elements in an appropriate underlying lumped model. For example, with a specification of 5% error, transmission-line characteristic impedance or velocity of propagation can be off by up to 5%. Equivalently, the series inductance or shunt capacitance per unit length can be off by 5%.

3. Quantitative error determination

That's a nice error definition, but how do we measure it? Comparisons with measurements are difficult. Even if error bars were calculated for measured data, it is likely that validation going below several percent error would lack significance. What if we could have measurements of some structure that was exact? We don't have exact measurements, but we can use the next best thing, a structure for which there is an exact theoretical solution. In scattering, a sphere is ideal. However, we are working with three-dimensional planar circuits (the third dimension is vertical vias; we reserve 2.5D for structures with 3D fields, but no z-directed current). The structure we use is the stripline standard [2]. It is one of the few planar structures which has an exact solution. We found that error due to cell length causes error in velocity of propagation, while error in cell width causes error in characteristic impedance. An elegantly simple result, when you think about it. The error model (using roof-top basis functions) is

$$e_T \le \frac{16}{N_w} + 2\left(\frac{16}{N_L}\right)^2, \quad N_w > 3, \quad N_L > 15$$

where e_T is the total error, N_w is the number of cells across the line width, and N_L is the number of cells along the line per wavelength. If variable width cells are used, N_w should be based on the cell width at the line edge. Note that developing an error model in this detail would be nearly impossible, if we were to use measured data instead of an exact standard

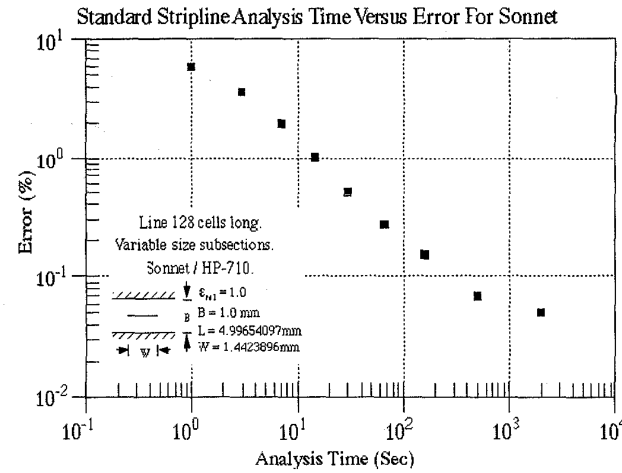

Figure 1. A performance plot for the *Sonnet* analysis shows uniform convergence to better than 0.1%. The final cell width corresponds to 512 cells across the width of the line.

This standard is also used [3] to generate "performance plots," as in Figure 1. Here, we compare analysis time versus error. Since the benchmark is exact, we can determine error to well below the 0.1% level, impossible without exact data. Wide ranges in performance, different rates of convergence, non-monotonic convergence, and even divergence [4] have been found in this way. Note that any direct comparisons of different techniques should be performed by an objective, independent, third party, with equal levels of experience in all techniques being compared.

Of course, no standard can evaluate all error sources. This standard's primary purpose is to evaluate error due to finite cell size. While there can be many error sources, cell-size error is, by definition, the principal error in an asymptotically exact analysis. The resulting cell-size error model is then used to decide the starting cell size for a complex circuit. While significantly larger cell sizes are capable of providing low error results (see the next section), such situations are special cases, and entail significant risk.

4. Error cancellation and software validation

In the course of investigating the stripline standard, we found that when cell width is comparable to cell length, length error can cancel width error. We quickly found that this error cancellation is unusable in a design situation, because it is highly dependent on the specific circuit and cell sizes. We did find, however, that the cell sizes used for most published microwave-software validations are in the prime range for error cancellation to occur. We expect that this is accidental in research situations. However, error cancellation can provide up to a ±5% tuning range, adjusted by simply modifying the cell-size aspect ratio. Because of this, a good validation must always be a "double blind." In other words, the provider of the validation structure must not know the analysis results in advance, and, most importantly, the provider of the analysis must not know the measured results in advance. Both should do their work separately, and then compare results after each is done.

However, the only way to be sure a calculated result has not been "tuned," or that error cancellation has not been accidentally invoked by a lucky choice of cell dimensions, is to perform a convergence analysis, showing that the analysis converges uniformly to the correct result. Given an asymptotically exact analysis and, in our case, two cell dimensions, two such analyses are usually required. The first sets the cell size along one dimension to a very small size, and varies the other. The second does the same with the other cell dimension. It is a mistake to vary both cell dimensions simultaneously, as the two error sources can "beat" against each other, yielding a non-monotonic convergence, which can not be extrapolated to the exact result.

Presently, microwave-software validations rarely present convergence analyses. They are often considered to be of no interest. In view of the error cancellation possible when using the large cell dimensions usually published, convergence analyses (like the performance plot of Figure 1) really need to be included in published validations. A single analysis, with the usual warm, fuzzy statement of "excellent agreement," just isn't good enough anymore..

It is sometimes suggested that we develop more complex standards. The principal difficulty here is to find a structure sensitive to analysis error, which can be fabricated and measured to better than 1% total error. If there is more error, differences between measured and calculated results can be attributed to measurement error, and the ubiquitous "excellent agreement" is then reported. In the author's opinion, when a validation experiment reports "excellent agreement," the experiment has failed, not succeeded. Specifically, it has failed to quantitatively measure the analysis error. Complex standards, fabricated and measured with a documented ultra-low error, are a difficult and much needed area of research, which is not currently being addressed.

5. Design validation

The stripline standard-error model gives a good starting point for a cell size likely to achieve a desired error performance. But what should an engineer do when the result absolutely, positively must be within the required specifications the first time? The answer here is also a convergence analysis. As shown in the next example, once we have confidence that an analysis is asymptotically exact, the engineer can evaluate successively smaller cell sizes, until convergence to within the design specifications is explicitly demonstrated. If additional accuracy is needed and the convergence is strong and monotonic, the engineer can simply extrapolate the convergence results.

As an example of a design validation, we recently analyzed a three-resonator 1%-bandwidth superconducting filter. Measured data were not available at the time (this was a double-blind test); however, results were needed to ±0.1% for center frequency. We performed a series of analyses, until we could demonstrate that the desired error level had been achieved (Figure 2). This was confirmed by subsequent measurement.

In the above case, if we had relied on an initial "lucky guess" for cell size, we might have chosen a cell size small enough, or a cell size which invoked error canceling, but we would not have been sure. The convergence analysis explicitly demonstrates that we have met the design requirements, without a priori access to the measured data, thus eliminating unnecessary risk.

We predict that in the next several years, convergence analysis will become a standard part of the microwave-design cycle, for

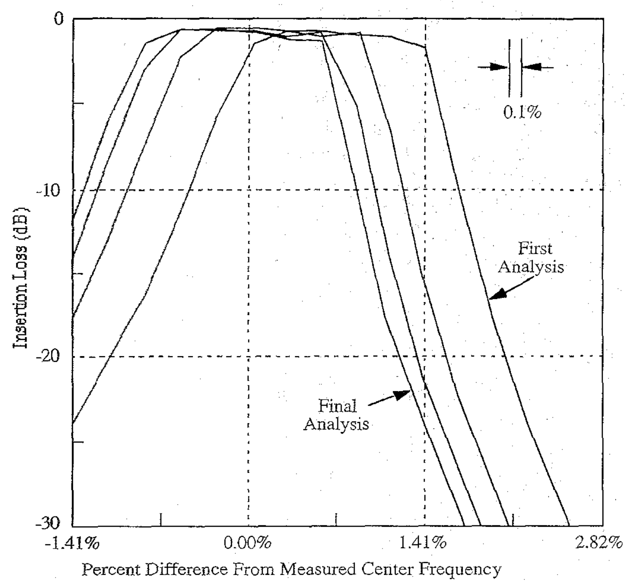

Figure 2. A convergence analysis for a three-resonator 1%-bandwidth filter. Each successive analysis uses cells one-half the length of the previous analysis. In this way, the required ±0.1% error is verified. Measured data (not presently available for publication) confirms the result.

the simple reason that it is much cheaper to do a second analysis than it is to do a second design and fabrication.

6. Publishing validation

Publishing papers on quantitative-error validation is difficult. The basic problem, in my view, is that papers on analysis error and software validation tend to be applications oriented, of primary interest to the applied microwave designer who faces difficult specifications on a daily basis. In my experience in the microwave field, research-oriented reviewers do not face such situations, and tend to view error-analysis research with little interest. If this situation is to change, then applications-oriented reviewers must be brought into the process. After all, we would not consider having an applications reviewer review a theoretical paper. Why should the reverse be allowed?

There is also considerable effort with regard to error analysis in nearly all non-microwave fields of science and engineering. How can we learn from all that is going on? I'm sure that the efforts of the Editor of this column to bring in outside authors is at least a part of the answer. But even if we can get outside researchers interested, and get our own researchers working on error analysis, it will all be for nothing if they can not publish because reviewers consider such material to be of little interest.

In our work, we find a tremendous thirst for hard information on analysis error among microwave-design engineers. While difficulty in publishing can be expected, I assure any researcher that good, quantitative work in this area is of tremendous utility to the working engineer. If you want to do genuinely useful research, this is an area with real need.

Lack of interest in the research community is certainly part of the problem here in the microwave world. But there are enough supporters of these concepts that we should eventually get a lively discussion going, much along the lines of what you have in the antenna community today.

7. References

1. J. C. Rautio, "Experimental validation of electromagnetic software," *International Journal of Microwave and Millimeter-Wave Computer-Aided Engineering*, **1**, 4, October 1991, pp. 379-385.

2. J. C. Rautio, "An ultra-high precision benchmark for validation of planar electromagnetic analyses," *IEEE Transactions on Microwave Theory and Techniques*, **MTT-42**, 11, November 1994, pp. 2046-2050.

3. J. C. Rautio, "MIC simulation column–a standard stripline benchmark," *International Journal of Microwave and Millimeter-Wave Computer-Aided Engineering*, **4**, 2, April 1994, pp. 209-212.

4. J. C. Rautio, "MIC simulation column," *International Journal of Microwave and Millimeter-Wave Computer-Aided Engineering*, **5**, 5, September 1995, pp. 365-367.

Error Estimation for Numerical Differential Equations

Donald J. Estep[1], Sjoerd M. Verduyn Lunel[2], and Roy D. Williams[3]

[1]School of Mathematics
Georgia Institute of Technology
Atlanta, GA 30332, USA
E-mail: estep@math.gatech.edu

[2]Faculteit Wiskunde en Informatica
Universiteit van Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam
Netherlands
E-mail: verduyn@fwi.uva.nl

[3]Center for Advanced Computing Research
California Institute of Technology
Pasadena, CA 91125, USA
E-mail: roy@cacr.caltech.edu

Given a capable human being and a computer, it is possible to make an approximation to the solution of a nonlinear differential equation. However, under the (usually correct) assumption that the equation is analytically intractable, the result of the computation is not the exact solution; indeed it may be so far from the exact solution as to be completely useless. We are interested in the relationship between the effort expended by the human and the computer, and the quality of the computed approximation to a partial or ordinary differential equation. To be specific, we would like to think in terms of a cost-benefit analysis. The cost of the computation is a combination of the human effort and computer resources used to obtain the approximation. The benefit includes, of course, the computed approximation, but it also includes an estimate of the quality of the approximation, that is, an error estimate. It is our opinion that in computational science, as with the experimental sciences, results should always be presented with some estimate of their accuracy. In addition, however, there is another facet to error estimation: one cannot even attempt a cost-benefit analysis or efficiency comparison of methods without an error estimate to evaluate the results.

Let $y(t)$ be the exact solution of a partial or ordinary differential equation

$$\frac{\partial y}{\partial t} + f(y,t) = 0,$$

with suitable initial and boundary conditions, where f may involve space derivatives. We compute a numerical approximation, Y, with respect to a method-of-lines discretization of the space-time domain, and we use $h(x,t)$ and $k(t)$, respectively, to denote the size of the space and time mesh at position (x,t). We emphasize that these may vary at different points in time and space. For example, the time step from time t to $t+k$ may change from one step to the next. We compute an approximation by choosing a function from a finite-dimensional space (such as polynomials) that approximately satisfies the differential equation in both space and time.

For example, the numerical solution might satisfy an approximation of the differential equation in the form of a difference scheme, or it might be chosen to be the best solution of the equation among all functions in the finite-dimensional space in a finite-element method. The error is $e = y - Y$, and this article is about estimating its norm, $\|e\|$. The norm may be chosen to emphasize error at some particular time and point in space, or an average over the whole space-time domain, or something else.

In the classical a-priori error theory, we attempt to estimate the error before computing the approximation, by first bounding the error produced in a step, and then assuming that errors accumulate in the worst possible way, to bound the global error. We measure the error of the approximation over a single time step as an interpolation error, I, that is bounded using Taylor's theorem

$$I \le C_n \left\| h^n D^n y \right\| + C_m k^m \left\| D^m y \right\|,$$

where D_n and D_m denote a combination of space and time derivatives of the appropriate order for the accuracy, C_n and C_m are constants that depend on the method for computing Y, and the norms are localized to the space domain and the current time step.

We then have to bound the effects of accumulation of such errors over all the subsequent intervals, in order to bound the total error. We can analyze the effect of perturbations in a solution of a nonlinear problem by studying the differential equation obtained by linearizing around the solution. The coefficient matrix for the linearized problem is the Jacobian matrix $\partial f / \partial y$. If we write J_{max} as the maximum norm of the Jacobian over "all possible values of y," we can get a global error bound: