QUANTIFICATION OF UNCERTAINTY IN COMPUTATIONAL FLUID DYNAMICS

P. J. Roache

Ecodynamics Research Associates, Inc., P.O. Box 9229, Albuquerque, New Mexico 87119

ABSTRACT

This review covers Verification, Validation, Confirmation and related subjects for computational fluid dynamics (CFD), including error taxonomies, error estimation and banding, convergence rates, surrogate estimators, nonlinear dynamics, and error estimation for grid adaptation vs Quantification of Uncertainty.

Introduction

The focus of this review is the *quantification of uncertainty*, the estimation or *banding* of the numerical error of a final calculation in computational fluid dynamics (CFD) and related fields in computational physics, mathematics, and engineering. By *final calculation* I mean one considered to be used as is. The motivation for this calculation is different than that for an error estimate used for the process of solution-adaptive grid generation. Although the present methods may be applicable to grid adaptation, and some of the methods described herein were motivated by that problem, the grid adaptation problem has vastly different (though usually unacknowledged) requirements.

The key word is *quantification* of uncertainty, as opposed to vague and all too common qualitative assessments. Quantification of uncertainty may also involve more than just obtaining a good error estimate; in fact, the more conservative procedure based on the grid-convergence index (described below) reports an error band equal to three times an error estimate. Furthermore, I consider

123

here only *a posteriori* error estimation, being of the opinion that useful *a priori* estimation is not possible for nontrivial fluid mechanics problems.

Policy Statements on Numerical Uncertainty

In 1986, the editors of the *ASME Journal of Fluids Engineering* (*JFE*) saw fit to publish a brief policy statement (Roache et al 1986) requiring at least minimal attention to the quantification[1] of numerical accuracy. Although the statement seemed innocuous and seemed to address an obvious need, it still met significant resistance. (Experience with the implementation of the policy is given in Roache 1990.) Since then, other journals have adopted similar explicit policies (ASME Editorial Board 1994, AIAA 1994, Gresho & Taylor 1994). The *JFE* has expanded its original policy, including a prohibition of methods with first-order spatial accuracy (Freitas 1993a, 1995b; see discussion in Shyy & Shindir 1994, Vanka 1994, Roache 1994b, Freitas 1994; Leonard 1995). An associate editor of the *AIAA Journal* (W Oberkampf 1995, personal communication) estimates that he needs to cite the journal policy in approximately three quarters of his communications with manuscript authors.

The topic has also been the subject of several American Society of Mechanical Engineers (ASME) symposia and resulting series of special publications (Celik & Freitas 1990, Celik et al 1993, Johnson & Hughes 1995), from which the present article takes its title. Also see the related ASME publications on the "CFD Triathlons" (Freitas 1993b, 1995a) and other benchmark comparison exercises, for example, the international HYDROCOIN and PSACOIN projects (OECD 1988, 1990). These exercises give a fair cross-section of present CFD practices in regard to numerical accuracy and generally are far from edifying. Oldenburg & Pruess (1995) give an example of qualitatively different flow revealed by even a cursory grid resolution study that was not discovered by any participants in the HYDROCOIN code intercomparison exercise.

Verification and Validation: Numerical versus Conceptual Modeling

It is useful at the onset to make the important semantic distinction between *verification* and *validation*. Following Boehm (1981) and Blottner (1990), I adopt the succinct description of verification as "solving the equations right," and of validation as "solving the right equations." The code author defines precisely what partial differential equations (PDEs) are being solved and demonstrates convincingly that they are solved correctly—that is, usually with some

---

[1]Other aspects of *confidence building* in CFD, including broader issues of code quality-assurance, confirmation, calibration, tuning, and certification, are found in Mehta (1989, 1991, 1995), Roache et al (1990), Aeschliman et al (1995), Oberkampf (1994), Oberkampf et al (1995), Cosner (1995), Melnik et al (1995), and Roache (1995).

order of accuracy and always consistently, so that as some measure of discretization Δ (e.g. the mesh increments) approaches zero, the code produces a solution to the continuum PDEs; this is verification. Whether or not those equations and that solution bear any relation to a physical problem of interest to the code user is the subject of validation. In a meaningful sense, a code cannot be validated; only a calculation (or range of calculations with a code) can be validated. In my experience, dealing with other than algorithm developers, this is a difficult concept and requires frequent reiteration.

Another way to make the distinction (i.e. to get to the idea behind the words, beyond "mere" semantics) is to speak of numerical errors versus conceptual modeling errors. An example of conceptual modeling versus numerical modeling is the assumption of incompressibility. This is clearly a conceptual modeling assumption. Is it the code builder's fault, or any criticism of the code itself, if the user incorrectly applies it? For example, dynamic stall involves compressibility at a surprisingly low free-stream Mach number. Results from an incompressible code may not agree with experiment very well, but we cannot say that the code fails verification because it was applied to a compressible flow problem, although we may have some sympathy for the user who is fooled by dynamic stall. But no one would have sympathy for a user who applied an incompressible flow code to a reentry vehicle at Mach 20. In such a case, the lack of agreement with experiment is not a code problem but a modeling problem. The same is true of many practical aspects of applying a CFD code. The model includes more than the code. "Model" includes conceptual modeling assumptions (e.g. incompressibility, symmetry, etc). It also includes data input to the code—e.g. geometry data, which often are not so easy to determine accurately as one might expect, and boundary conditions and initial conditions. Incorrect determination of any of these items can lead to failure of validation of a model, with possibly no criticism of the code.

The question of solution uniqueness always arises with nonlinear equations, and its position in the verification-validation distinction is important. If, as stated above, we take verification to mean simply that *a* solution to the continuum PDEs is obtained, then the problem of PDE nonuniqueness is avoided. This is probably being too easy on the code developer in some cases. Physically inadmissable solutions should be eliminated by the code—for example, a shock-tube simulator should be required to eliminate expansion shock solutions; but nature abounds with physically nonunique solutions in fluid dynamics—for example, hysteresis of airfoil stall and recovery, and bi-stable fluid amplifiers. If nature cannot decide which solution to produce, we cannot expect more of a code. Note that obtaining "a solution to the continuum PDEs" may involve sorting through multiple numerical solutions for nonlinear problems; for example, see later discussion of Yee et al (1991) and Lafon & Yee (1992). See also

especially Stephens & Shubin (1981), whose study of Euler solutions indicates that the multiple numerical solutions converge to the same solution as the grid is refined.

A related question is the stability of a solution. Although people usually think of artificial (i.e. numerical artifact) instability as a code difficulty, the equally common situation is *artificial stability*, for example, a first-order time-dependent code whose artificial viscosity damps out disturbances so that accurate calculations of physical fluid instability would require impractically high resolution. In my opinion, it is asking too much of a code to exactly mimic stability boundaries except in the limit of $\Delta \to 0$. For example, it is not a failure of verification for a full Navier-Stokes code to produce steady shear-layer solutions at Reynolds numbers known to be unstable; an unstable solution is still a valid solution.

The choice of using verification or validation was originally arbitrary and is now recommended solely because of common developing use. (I have published articles using the opposite definition.) In a common English thesaurus, the terms *verify*, *validate*, and *confirm* are all synonyms, but the words are used herein, and generally in code quality-assurance (QA), as technical terms with more context-specific meaning. Such technical terms are preferably related to common use, but each term's technical meaning is defined independent of common use and in a specific technical context. This is not a universally accepted attitude toward semantics. In a widely quoted paper that has been recently described as brilliant in an otherwise excellent *Scientific American* article (Horgan 1995), Oreskes et al (1994) think that we can find the real meaning of a technical term by inquiring about its common meaning. They make much of supposed intrinsic meaning in the words *verify* and *validate* and, as in a Greek morality play, agonize over truth. They come to the remarkable conclusion that it is impossible to verify or validate a numerical model of a natural system. Now most of their concern is with groundwater flow codes, and indeed, in geophysics problems, validation is very difficult. But they extend this to all physical sciences. They clearly have no intuitive concept of error tolerance, or of range of applicability, or of common sense. My impression is that they, like most lay readers, actually think Newton's law of gravity was proven wrong by Einstein, rather than that Einstein defined the limits of applicability of Newton. But Oreskes et al (1994) go much further, quoting with approval (in their footnote 36) various modern philosophers who question not only whether we can prove any hypothesis true, but also "whether we can in fact prove a hypothesis false." They are talking about physical laws—not just codes but any physical law. Specifically, we can neither validate nor invalidate Newton's Law of Gravity. (What shall we do? No hazardous waste

disposals, no bridges, no airplanes, no . . . .) See also Konikow & Bredehoeft (1992) and a rebuttal discussion by Leijnse & Hassanizadeh (1994). Clearly, we are not interested in such worthless semantics and effete philosophizing, but in practical definitions, applied in the context of engineering and science accuracy.

Inadequate Error Taxonomies

Several taxonomies of errors given in the literature are inadequate and misleading, in my opinion. Not all lists are taxonomies. For example, the list "flora, fauna, mammals, dogs" is not an adequate biological taxonomy. Mammals are not separate from fauna but are part of it, as are dogs. But we see this kind of false taxonomy often in CFD.

For example, so-called *grid-generation errors* (Ferziger 1993) are not separate from discretization errors. For the verification of a code or a calculation, there are no such things as grid-generation errors [nor are there "errors associated with coordinate transformations" (Ferziger 1993)]. Indeed, bad grids add to discretization error size, but they do not add new terms. This does not mean that one grid is as good as another, or that a really bad grid cannot magnify errors, but only that these so-called grid-generation errors do not need to be considered separately from other discretization errors in a grid-convergence test. If the grid-convergence test is performed, and the errors are shown to reduce as $O(\Delta^2)$, for example, then all discretization errors are verified. One does not need to separately estimate or band the grid-generation errors.

Likewise for Karniadakis's (1995) proposed numerical error bar (see also Vanka 1995), which consists of separately estimated numerical errors from *boundary conditions, computational domain size, temporal errors*, and *spatial errors*. This is clearly a false taxonomy. Numerical errors at boundaries can be ordered in Δ [e.g. the various approaches for wall vorticity (Roache 1972)] or possibly nonordered in Δ [e.g. the boundary-layer-like $\partial p / \partial n = 0$, or the downstream (outflow) boundary location (Roache 1972)]. Ordered errors will tend to zero as the discretization improves, so that a boundary error from wall vorticity evaluation need not be considered separately from spatial errors. That taxonomy (Karniadakis 1995) already includes temporal errors and spatial errors and computational domain size errors, so that both ordered and nonordered boundary errors are already counted elsewhere in the taxonomy. Since the intention in (Karniadakis 1995) is to provide a quantitative breakdown in the sources of numerical error in an error band, the proposed taxonomy is not only confused but misleading.

Note, however, that outflow boundary errors may prove to be ordered not in Δ but in $1/l$, where l is the distance from the region of interest (e.g. an airfoil)

to the outflow boundary. [See Zingg's (1993) data, shown in Roache (1994a) to be first-order in $1/l$.]

The subject of outflow boundary conditions does produce some fuzziness in categorization of Verification vs. Validation, in my opinion. The error can be ordered, as above, and therefore can be part of verification. That is, it is up to the user (who is doing the conceptual modeling) to estimate or band the error caused by the position of the outflow boundary. But if the code has some sophisticated outflow condition (e.g. a simple vortex condition for Euler equations), then the distinction is not so clear. Certainly the equations used are clear, and the code may "solve the equations right" (i.e. verification), yet there exists another benchmark solely from the mathematics (the case with infinite boundary distance) which could be used to justify the outflow condition without recourse to physical experiment (which would be associated with validation).

Another example of the same semantic failure arises when we consider benchmarking a turbulent boundary-layer code or parabolized Navier-Stokes (PNS) code against a Reynolds-stress-averaged full Navier-Stokes (NS) code. Presume that both codes are convincingly verified, that is, they correctly solve their respective equations. Suppose that the PNS code results agree well with the NS code results for some range of parameters (e.g. including angle of attack). This agreement is not included in the term *verification*, since the verification of the PNS code has already been completed prior to the NS benchmarking. Then we could say that the agreement has demonstrated that the PNS code is "solving the right equations" in one sense, that is, it justifies the use of parabolic marching equations. Yet to claim validation would be over-reaching, since we have not demonstrated the adequacy of the turbulence model by comparison with experiment. We have "solved the right equations" only in an intermediate sense of demonstrating that the PNS equations adequately represent the full NS equations, but not in the ultimate sense of validation, of "solving the right *physical* equations."

Unfortunately, such "mere semantics" may become of vital interest to code QA when dealing with regulatory agencies such as the EPA or with legal definitions in a NASA contract.

Taxonomy for Additional Information for Error Estimates

Once we have produced a CFD solution of the governing partial differential equations, it is clear that we require some additional information in order to quantitatively estimate the uncertainty or numerical accuracy. The following taxonomy of sources of this additional information provides a framework for the discussion. By the word *grid*, I refer to any measure of discretization (e.g. Cartesian grid, nonorthogonal grid, number of Fourier modes in a spectral solution, number of discrete vortices).

Sources of Additional Information for Error Estimation, Given a CFD Solution of the Governing PDEs on a Grid

A. Additional Solution(s) of the Governing Equations on Other Grids

- A.1 Grid Refinement
- A.2 Grid Coarsening
- A.3 Other Unrelated Grid(s)

B. Additional Solution(s) of the Governing Equations on the Same Grid

- B.1 Higher-order Accuracy Solution(s)
- B.2 Lower-order Accuracy Solution(s)

C. Auxiliary PDE Solutions on the Same Grid

D. Auxiliary Algebraic Evaluations on the Same Grid; Surrogate Estimators

- D.1 Nonconservation of Conservation Variables
- D.2 Nonconservation of Higher Moments
- D.3 Zhu-Zienkiewicz-type Estimators
- D.4 Convergence of Higher-order Quadratures

The following are brief remarks on this taxonomy, some of which will be justified and amplified in later discussion:

Categories A and B involve the direct, unambiguous evaluation of any error measure of engineering or scientific interest. For Category A, no additional CFD code development or modifications are required. For Category B, no additional grid generation is required.

The approach of Category C (e.g. Van Straalen et al 1995) does not simply involve a local evaluation of something. The key aspect here is that errors are transported (i.e. advected, diffused). However, a simple local evaluation of something, without advection, is just what you want to guide solution adaptation, hence, the different needs of error estimation for solution adaptation versus quantification of uncertainty for a final calculation.

Category D, involving algebraic evaluations on the same grid, is cheap, needs no additional grid generation, and uses no significant dynamic memory. However, the error evaluated usually has no direct relation to any error measure of engineering or scientific interest.

Grid Refinement and Coarsening

Systematic grid-convergence studies are the most common, most straightforward and arguably constitute the most reliable technique for the quantification

of numerical uncertainty. Unlike the other methods available, this approach can be used to dependably consider the convergence of any quantity of interest, as well as the usual L_2 and L_∞ norms.

By *grid-convergence studies* people usually mean Category A.1 (grid refinement), but Category A.2 (grid coarsening) usually would make more sense. If completely solved solutions are obtained on two grids, presumably the finer grid solution would be used, so the coarse grid solution could be used to estimate the error. Whether one refines or coarsens just depends on which grid was calculated first. So for completely solved solutions, grid refinement and coarsening are identical.

A disadvantage of Category A methods is that multiple-grid generation is required. Cartesian grids obviously pose no problem. For boundary-fitted structured grids, the simplest method for grid doubling (halving) is to generate the finest grid first, using whatever method is preferred (e.g. see Knupp & Steinberg 1994) and then obtain the coarser grids by removing every other point (e.g. see Zingg 1993). For noninteger grid refinement (coarsening), the same generating equations and parameters should be used. See Roache (1993a, 1994) for further discussion, including nonstructured grid refinement (see also Pelletier & Ignat 1995) and structured refinement of nonstructured grids.

Error estimation using other unrelated grids (Category A.3) poses an interesting challenge. By *unrelated grids* I mean two or more grids that are overlapping but not simply obtained one from the other. For example, say grid A is finer than grid B in some regions but coarser in others, such as might be obtained in two steps of an r-type (redistribution) solution-adaptive grid. It would seem that the two solutions on unrelated grids would provide the additional information necessary to estimate the uncertainty in either, but to my knowledge a method for doing so has not been invented.

In order to quantify the uncertainty with systematically refined (coarsened) grids, we need the convergence rate p to estimate the error. For now, we are assuming that p is known, that is, that we are using a rigorously verified code and are now concerned with quantification of the uncertainty of a particular calculation using two grid solutions. The same methods can be converted to verify a code, that is, to verify (or determine) p. Code verification is discussed in a later section.

A Method for Uniform Reporting of Grid-Convergence Studies: The Grid-Convergence Index

This section presents a summary of the main results of Roache (1993a,b; 1994a). The grid-convergence index (GCI) presents a simple method for uniform reporting of grid-convergence studies without any restriction to integer refinement

(e.g. grid doubling). The GCI is based on generalized Richardson Extrapolation involving comparison of discrete solutions at two different grid spacings.

Richardson Extrapolation

Richardson Extrapolation, also known as h^2 extrapolation, the deferred approach to the limit, and iterated extrapolation, was first used by Richardson in 1910 and later embellished in 1927. The discrete solutions f are assumed to have a series representation, in the grid spacing h, of

$$f = f[exact] + g_1 h + g_2 h^2 + g_3 h^3 + \cdots \tag{1}$$

The functions g_1, g_2, and so on are defined in the continuum and do not depend on any discretization. For infinitely differentiable solutions, these functions are related to all orders to the solution derivatives through the elementary Taylor series expansions, but this is not a necessary assumption for Richardson Extrapolation, nor is the infinite series indicated in Equation 1. It is only necessary that Equation 1 be a valid definition for the order of the discretization. Thus, the extrapolation may be valid for finite element solutions, and so on.

For a second-order method, $g_1 = 0$. Then the idea is to combine two separate discrete solutions, f_1 and f_2, on two different grids with (uniform) discrete spacings of h_1 (fine grid) and h_2 (coarse grid), so as to eliminate the leading-order error terms in the assumed error expansion, that is, to solve for g_2 at the grid points in Equation 1, substitute this value into Equation 1, and obtain a more accurate estimate of $f[exact]$. The result is the original statement (Richardson 1927) for h^2 extrapolation:

$$f = f[exact] + \left(h_2^2 f_1 - h_1^2 f_2 \right) / \left(h_2^2 - h_1^2 \right) + \text{ H.O.T.} \tag{2}$$

where H.O.T. are higher-order terms. Using the grid refinement ratio $r = h_2 / h_1$, this result can be conveniently expressed in terms of a correction to the fine-grid solution f_1, dropping H.O.T.:

$$f[exact] \cong f_1 + (f_1 - f_2)/(r^2 - 1). \tag{3}$$

The most common use of this method is with a grid doubling, or halving. (As noted earlier, these are identical, that is, we just have a coarse grid and a fine grid.) With $r = 2$, Equation 3 becomes

$$f[exact] \cong 4/3 f_1 - 1/3 f_2. \tag{4}$$

It is often stated that Equation 4 is fourth-order accurate if f_1 and f_2 are second-order accurate. Actually, as known by Richardson, this statement is true only if odd powers are absent in the expansion (Equation 1), which he achieved by assuming the exclusive use of second-order centered differences. If uncentered differences are used (e.g. upstream weighting of advection terms),

even if these are second-order accurate (3-point upstream), the h^2 extrapolation is third-order accurate, not fourth. As a practical limitation, even extrapolations based on centered differences do not display the anticipated fourth-order accuracy until the cell Reynolds number Rc is reduced; for the 1-D advection-diffusion equation with Dirichlet boundary conditions, $Rc < 3$ is required (Roache & Knupp 1993).

Although Richardson Extrapolation is most commonly applied to grid doubling and is often stated to be applicable only to integer grid refinement (e.g. Conte & DeBoor 1965), this is not required. In order to use Equation 3, it is necessary to have values of f_1 and f_2 at the same points, which would seem to require commonality of the discrete solutions and, therefore, integer grid refinement ratios r (grid doubling, tripling, and so on). However, even in his 1910 paper, Richardson looked forward to defining a continuum f_2 by higher-order interpolation, and in his 1927 paper had a specific approach worked out. Ferziger (1993) alludes to this approach with less detail but more generality. Similarly, Richardson Extrapolation is commonly applied only to obtaining a higher-order estimate on the coarse grid with $h_2 = 2h_1$, but Roache & Knupp (1993) show how to obtain fourth-order accuracy on all fine-grid points by simple second-order interpolation, not of the solution values f_2, but of the extrapolated correction from Equation 4, that is, by second-order interpolation of $1/3(f_1 - f_2)$. The use of simple second-order interpolation avoids complexities with nonuniform grids and near-boundary points.

Richardson (1910, 1927) also considered sixth-order extrapolation (using three grid solutions to eliminate g_2 and g_4), parabolic and elliptic equations, staggered grids (then called interpenetrating lattices), rapid oscillations and the $2h$ wavelength limit, *a priori* error estimates, singularities, integral equations, statistical problems, Fourier coefficients, and other noncalculus problems. For example, Richardson (1927) showed the power of the method in an elegant example of extrapolating two very crude approximations to a circle, namely an inscribed square and an inscribed hexagon, to get an estimate of π with three-figure accuracy.

The usual assumptions of smoothness apply, as well as the assumption (often verified) that the local error order is indicative of the global error order. The extrapolation must be used with considerable caution, since it involves the additional assumption of monotone truncation-error convergence in the mesh spacing h. This assumption may not be valid for coarse grids. Also, the extrapolation magnifies machine round-off errors and incomplete iteration errors (Roache 1972). In spite of these caveats, the method is extremely convenient to use compared to forming and solving direct fourth-order discretizations, which involve more complicated stencils, wider bandwidth matrices, special considerations for near-boundary points and non-Dirichlet boundary conditions,

additional stability analyses, and so on, especially in nonorthogonal coordinates that generate cross-derivative terms and generally complicated equations. Such an application was given by the present author in Roache (1982).

The method is in fact oblivious to the equations being discretized and to the dimensionality of the problem and can easily be applied as a postprocessor (Roache 1982) to solutions on two grids with no reference to the codes, algorithms, or governing equations that produced the solutions, as long as the original solutions are indeed second-order accurate. (By this somewhat abusive but common terminology, I mean a solution obtained with a second-order accurate method applied in its asymptotic range.) The difference between the second-order solution and the extrapolated fourth-order solution is itself a useful diagnostic tool, obviously being an error estimator (although it does not provide a true *bound* on the error except possibly for certain trivial problems). It was used very carefully, with an experimental determination rather than an assumption of the local order of convergence, by de Vahl Davis (1983) in his classic benchmark study of a model free-convection problem. See Nguyen & Maclaine-Cross (1988) for application to heat exchanger pressure drop coefficients. Zingg (1993) applied the Richardson error estimator to airfoil lift and drag calculations in body-fitted grids. Zingg's work demonstrates the necessity of grid-convergence testing even when experimental data are available. In four of seven cases, experimental agreement was better with coarse-grid calculations than with fine! Also, his data indicate that Richardson Extrapolation can be applied to the estimation of far-field boundary errors, with the error being first order in the inverse of distance to the boundary (Roache 1994a). Blottner (1990) has used the same procedure to estimate effects of artificial dissipation terms in hypersonic flow calculations.

An important aspect of Richardson Extrapolation is that it applies not only to point-by-point solution values but also to solution functionals (e.g. lift coefficient C_L for an aerodynamics problem or integrated discharge for a groundwater flow problem), provided that consistent or higher-order methods are used in the evaluations (e.g. second- or higher-order quadratures for lift) as well as the basic assumption that the order of the method applies globally as well as locally. If Richardson Extrapolation is applied to produce (say) fourth-order accurate grid values, one could in principle calculate a fourth-order accurate functional like C_L from the grid values, but it would require careful implementation of fourth-order accurate quadratures. It is much simpler to apply the extrapolation directly to the C_Ls obtained in each grid, requiring only second-order quadratures. Indeed, this is a major attraction of Richardson Extrapolation compared to using fourth-order accurate stencils solved either directly or by deferred corrections. (Note, however, that the two approaches yield different answers, although both are fourth-order accurate if done properly.)

A significant yet overlooked disadvantage of Richardson Extrapolation is that the extrapolated solution generally is not *conservative* in the sense of maintaining conservation properties (e.g. Roache 1972). This shortcoming could well dictate that Richardson Extrapolation not be used. For example, if it were used on the groundwater flow simulations for the Waste Isolation Pilot Plant (WIPP PA Dept. 1992), it would be more accurate in some norm but would introduce additional nonconservative (i.e. lack of conservation property) source terms into the radionuclide transport equation. de Vahl Davis (1983, p. 256) pointed out the more fundamental problem that the extrapolated solution is "no longer internally consistent; the values of all the variables do not satisfy a system of finite difference approximations." It is also noteworthy that Richardson (1927) pointed out that the accuracy of the extrapolation does not apply to arbitrarily high derivatives of the solution. The extrapolation can introduce noise to the solution, which, although low level, may decrease the accuracy of the solution higher derivatives.

Thus, it is not advocated (Roache 1993a,b, 1994a) that Richardson Extrapolation necessarily be used to improve the reported solution, since that decision involves these considerations and possibly others. What is advocated is that, regardless of whether Richardson Extrapolation is used to improve the solution, the GCI based on the generalized theory of Richardson Extrapolation should be used to uniformly report grid-convergence tests.

A fine-grid Richardson error estimator approximates the error in a fine-grid solution, f_1, by comparing this solution to that of a coarse grid, f_2, and is defined as

$$E_1^{fine} = \frac{\varepsilon}{1 - r^p}, \tag{5}$$

while a coarse-grid Richardson error estimator approximates the error in a coarse-grid solution, f_2, by comparing the solution to that of a fine grid, f_1, and is defined as

$$E_2^{coarse} = \frac{r^p \varepsilon}{1 - r^p}, \tag{6}$$

where

$$\varepsilon = f_2 - f_1, \tag{7}$$

f_2 = a coarse-grid numerical solution obtained with grid spacing h_2,

f_1 = a fine-grid numerical solution obtained with grid spacing h_1,

r = refinement factor between the coarse and fine grid

$\quad (r = h_2/h_1 > 1)$, and

p = formal order of accuracy of the algorithm.

The actual error A_1 of the fine-grid solution may be expressed (Roache 1993a, 1994a) as

$$A_1 = E_1 + \mathrm{O}\big(h^{p+l}, E_1^2\big), \tag{8}$$

where $l = 1$ generally or $l = 2$ if centered differences have been used. Thus, E_1 is an ordered estimator and is a good approximation when $E_1 \ll 1$.

As noted above, it is neither necessary nor often desirable to use $r = 2$, or grid doubling (halving). Accurate application of these generalized Richardson-based grid-error estimators requires that the observed convergence rate equals the formal convergence rate. This requirement implies that the leading-order truncation error term in the error series truly dominates the error.

To account for the uncertainty in these generalized Richardson-based error estimates due to various factors and to put all grid-convergence studies on the same basis as grid doubling with a second-order method, a safety factor is incorporated into these estimators and the GCI is defined for fine and coarse grids as

$$GCI_1^{fine} = F_s|E_1| \tag{9}$$

$$GCI_2^{coarse} = F_s|E_2| \tag{10}$$

$F_s > 1$ can be interpreted as a safety factor, since $F_s = 1$ gives $GCI = |E|$. That is, the error band reduces to the *best estimate* of the error, analogous to a 50% error band of experimental data. I recommended (Roache 1993a, b; 1994a) a more conservative value of $F_s = 3$. This value also has the advantage of relating any grid-convergence study (any r and p) to one with a grid doubling and a second-order method ($r = 2$, $p = 2$). I emphasize that the GCIs are not error estimators but are three (or F_s) times the error estimators, representing error bands in a loose statistical sense. Ostensibly, if we have a fine-grid and a coarse-grid solution, we would be expected to use the fine-grid solution, so reporting of the above fine-grid evaluation of GCI would apply. However, a practical scenario occurs for which the contrary situation applies, i.e. we use the coarse-grid solution. Consider a parametric study in which hundreds of variations are to be run. (For example, consider a 3-D time-dependent study of dynamic stall, with perhaps 3 Mach numbers, 6 Reynolds numbers, 6 airfoil-thickness ratios, 3 rotor-tip designs, and 2 turbulence models, for a total of 648 combinations.) A scrupulous approach would require a grid-refinement study for each case, but most engineers would be satisfied with one or a few good grid-refinement tests, expecting, for example, that a grid adequate for a NACA 0012 airfoil could be assumed to be adequate for a NACA 0015 airfoil. (In fact, this is often not justified by experience. For example, stall characteristics can be quite sensitive to thickness ratio.) So for the bulk of the stack of calculations,

we would be using the coarse-grid solution, and we want a GCI for this solution. That is, we derive a GCI from Equation 5, not as the correction to the fine-grid solution f_1, but as the correction to the coarse-grid solution f_2. In this case, the error estimate changes and must be less optimistic.

A recent application of the GCI to airfoil calculations is given by Lotz et al (1995), which demonstrates the power of the method without the need for integer grid refinement and its application with solution-adaptive grids. Pelletier & Ignat (1995) have shown that the GCI is applicable, at least in a rough sense, to unstructured grid refinement.

The motivation for using $F_s > 1$ is that $F_s = 1$ is analogous to a 50% error band on experimental data, which is not adequate. My originally recommended value (Roache 1993a,b; 1994a) of $F_s = 3$ is conservative and relates the grid convergence study to one with a grid doubling with a second-order method. For many reasons (see below) this is not unduly conservative when only two grids are used in the study. However, it is now clear that $F_s = 3$ is overly conservative for scrupulously performed grid convergence studies using three or more grid solutions to experimentally determine the observed order of convergence p (e.g. see the papers in Johnson & Hughes 1995). For such high-quality studies, a modest and more palatable value of $F_s = 1.25$ appears to be adequately conservative. However, for the more common two-grid study (often performed reluctantly, at the insistence of journal editors) I still recommend the value $F_s = 3$ for the sake of uniform reporting and adequate conservatism.

Code Verification

So far, this article has been concerned with error estimation, or verification of a calculation. It has been assumed throughout that the code itself has been verified independently. If a code has an error (not an ordered error, but a mistake), then performance of grid-convergence studies or other error-estimation techniques will not faithfully produce quantification of the uncertainty. However, the same techniques can be used to verify the code separately, including its order of convergence, by monitoring grid convergence toward a nontrivial exact solution.

It is always possible to obtain a nontrivial exact solution for this procedure, if necessary, by specifying the exact solution and adding the appropriate forcing function to the governing PDEs (Steinberg & Roache 1985, Shih et al 1989, Ethier & Steinmen 1994, Westerink & Roache 1995). The necessary condition is that the solution be nontrivial, that is, that it have significant solution structure to exercise higher-derivative calculations. Significantly, it is not at all necessary that the solution be realistic in any sense. Physically realistic solutions will be more convincing to the mathematically naive (and therefore have value in the real world of regulatory agencies, contract management, and so on), but the code verification is just as well accomplished by nonphysical solutions.

(For the alternative philosophy, see Oberkampf 1994, Oberkampf et al 1995, and Aeschliman et al 1995.) Additional realistic solutions (such as Sudicky & Frind 1982, corrected) can be used in additional *confirmation* exercises (Roache 1995).

I claim that this method of code verification via systematic grid-convergence testing (whether or not the GCI is used) is both general and rigorous, and has proven remarkably sensitive in revealing small and subtle coding errors and usage inconsistencies, as indicated by the following examples (Westerink & Roache 1995; Salari et al 1995):

A. In verification tests of a commercial groundwater flow code, a first-order error in a single corner cell in a strongly elliptic problem caused the observed convergence to be first-order accurate (Roache et al 1990).

B. In verification tests of our SECO_FLOW_2D variable density groundwater flow code, first-order extrapolation for ghost cell values of only one quantity (aquifer thickness) along one boundary caused the observed convergence to be first-order accurate (Roache 1993c,d).

C. In contaminant transport calculations (advection-diffusion + decay, retardation, and matrix diffusion), use of a plausible single-grid-block representation for a point source as the grid is refined introduces error in a finite volume (or block-centered finite difference) formulation. In this cell configuration, the cell faces align with the boundaries of the computational domain, and doubling the number of cells requires the location of the single cell representing the source to shift by $\Delta/2$. It is to be expected that the solution accuracy in the neighborhood of the source would be affected. But surprisingly, the accuracy of time-integrated discharge across boundaries far from the source was also degraded to first-order accuracy (Salari et al 1995).

D. The observed convergence rate of ostensibly second-order-accurate turbulent-boundary-layer codes (Wilcox 1993) can be degraded, apparently by conditional statements limiting eddy viscosity and defining the boundary-layer edge (DC Wilcox 1995, personal communication,).

E. Airfoil codes can exhibit the expected second-order convergence rates for lift and drag but less for moment, possibly because of approximations involved in applying quasi-periodicity across cut-planes of a C-grid (K Salari 1995, personal communication).

F. When a 2-D compressible driven cavity problem is solved with an even number of vertical grid lines, one must average velocity profiles from grid

lines on each side of the centerline rather than use only one; otherwise, the convergence will be first-order (Roache & Salari 1990).

Note that in Examples C and F apparently first-order behavior was obtained with second-order verified codes; the problem in Example C was a subtle conceptual modeling error and in Example F was a not-so-subtle postprocessing error. In both cases, the codes had no coding errors and were applied in the asymptotic range.

Esoteric Errors in Code Verification

General CFD or computational physics codes [more general than the simple Poisson equation in nonorthogonal coordinates of Steinberg & Roache (1985)] would be difficult to include in a theorem because of esoteric errors. The difficult aspects of the codes are not algebraic complexity [in Steinberg & Roache (1985) we convincingly verified 1800 lines of dense Fortran]; the more difficult and vexing problems come from option combinations and conditional differencing. Esoteric errors can arise because of nonlinear flux limiters like FCT, TVD, hybrid or type-dependent differencing, and so on (Steinberg & Roache 1985).

Code Verification versus Verification of Calculations

In this review, I distinguish between verification of codes and verification of individual calculations. Clearly, code verification must be done before the verification of a (real) calculation can proceed. However, this distinction works both ways; it is not enough to perform a calculation of a new problem with a verified code. A code may be rigorously verified to be (say) second-order accurate, but when applied to a new problem, this fact provides no estimate of accuracy or confidence interval, that is, of the size of the error (as opposed to the order of the error). Even though the real problem solution may be converging at the same rate (say second-order) as the code verification problem, it is still necessary to perform grid-convergence tests in order to band the numerical error. (It would be preferable to have different words for these two verification activities, but I am at a loss for a clarifying term.) The very important point, independent of the semantics, is that *use of a verified code is not enough*. This point is probably well recognized by present readers, but it is not universally so. Especially in the commercial CFD arena, user expectations are often that the purchase and use of a "really good code" will remove from the user the obligation of "doing his homework," that is, the straightforward but tedious work of verification of calculations via systematic grid-convergence studies. This unrealistic hope is sometimes encouraged by advertising.

Unfortunately, the situation may be even more difficult than this, in that a new calculation may not even exhibit the verified order of convergence. If the

code verification problem(s) are in some sense close to the real problem of interest, the order of convergence (e.g. second-order) from the code verification exercise may be assumed to hold for the real problem. However, there is a distinction to be made between the formal order of convergence (as indicated by analyses of the discretization), the *actual* asymptotic order of convergence, and the *observed* (or apparent) order of convergence. Note that there may be more than one formal analysis, especially for nonlinear problems, so that formal order of convergence is not necessarily unique. Then the actual asymptotic rate of convergence may differ from a formal rate, due to competition of truncation error terms for a particular problem (Westerink & Roache 1995). This situation typically arises when higher grid resolution reveals higher modes of the solution, which often occurs in atmospheric and ocean modeling (Westerink & Roache 1995, Dietrich et al 1990) and could be expected in high Reynolds number DNS (Direct Numerical Simulation) calculations. It is also possible, and not altogether unusual, to observe *superconvergence*, that is, an actual asymptotic convergence rate higher than a formal rate, due to cancellation (or partial cancellation) of space and time truncation errors. This observation is evident in the classical situation of (what are generally) first- or second-order methods for constant velocity advection problems producing the exact answers for unity Courant number (Roache 1972). Finally, the observed convergence rate p may differ from the actual asymptotic convergence rate due to failure to achieve the asymptotic range of grid resolution. This is especially obvious if solution singularities are present, in which case the method of grid refinement still can be used to estimate uncertainty, but local values of observed p near the singularity are required (de Vahl Davis 1983).

The cause of the discrepancy between an observed convergence rate and the actual asymptotic convergence rate may or may not be revealed by formal analysis. For example, in Roache & Knupp (1993), a method was presented for completing classical Richardson Extrapolation. The classical method combines two second-order solutions on a coarse grid and a doubled fine grid to produce a fourth-order accurate solution, but only on the coarse grid. The completed algorithm was devised to produce a fourth-order solution on the entire fine grid. However, initial grid-convergence studies indicated only a third-order convergence rate for the steady-state Burgers equation, even though the formal rate was fourth order. Further analysis by PM Knupp then showed that, because of truncation error competition particular to the steady Burgers solution, the higher-order terms in the error expansion include the Reynolds number parameter Re, and that the asymptotic range for fourth-order accuracy occurs only for Re $\Delta x \ll 3$.

For these reasons, it appears to be preferable to verify a code with a nonrealistic problem that has nontrivial but regular solution behavior (e.g. a *tanh*

function) rather than to use a realistic problem that displays the confusing truncation error competition.

Extraction of Convergence Rate from Grid-Convergence Tests

When using grid-convergence studies to estimate the size of the discretization error, it is necessary to have at least two grid solutions and an *a priori* knowledge of the convergence rate p, for example, to have verified that the code is second-order accurate ($p = 2$). If an exact solution is known or constructed (see above), it is straightforward to extract the order of convergence p from results of a systematic grid-convergence test using at least two grid solutions; this serves to verify a code. However, it is also desirable to verify the order p for an actual problem, since the observed order of convergence depends on achieving the asymptotic range, which is problem dependent, and the observed order may differ from the formal (theoretical) order, or from the order verified for a test case, for a variety of reasons (see discussion above and a more complete discussion in Westerink & Roache 1995). Without an exact solution for the actual problem, it is necessary to have at least three grid solutions to extract p. Thus, if there is any suspicion that the grid resolution is not in the asymptotic range, three grid solutions are necessary to verify (or determine) the observed rate of convergence and thereby allow estimation of the error.

Blottner (1990) and others use graphical means, plotting the error on log paper and extracting the order from the slope. This procedure requires evaluation of the error itself, which is of course generally not known. If the finest grid solution is taken to be the reference value (unfortunately, often called the *exact* value, which it obviously is not), then the observed order will be accurate only for those grids far from the finest, and the calculated order approaching the finest grid will be indeterminate. Blottner (1990) improves on this by estimating the exact value by Richardson Extrapolation (see also Shirazi & Truman 1989), but this procedure is somewhat ambiguous since the order is already needed in order to perform the Richardson Extrapolation.

If the grid refinement is performed with constant r (not necessarily $r = 2$, but constant), then the order can be extracted directly from three grid solutions, without a need for estimating the exact solution, following de Vahl Davis (1983, p. 254). With subscript 1 indicating the finest grid in the present notation,

$$p = \ln\left(\frac{f_3 - f_2}{f_2 - f_1}\right) / \ln(r). \tag{11}$$

A generalization of this procedure, not restricted to constant r, is possible using the generalized theory of Richardson Extrapolation (Roache 1993a, b; 1994a). For constant r, Equation 19 (typo corrected) of Roache (1994a) may be used to verify an assumed order p. (It is not necessary to use the GCI itself.) One

calculates

$$\alpha = GCI_{12}^{fine} / GCI_{23}^{fine}. \tag{12}$$

If $\alpha \approx r^p$, then p is the observed order. However, Equation 12 also requires r to be constant over the three grid set, and it cannot be used to calculate p directly since p is implicitly present in the GCIs. The more general procedure (Roache 1995a) is to solve the equation

$$\frac{\varepsilon_{23}}{(r_{23}^p - 1)} = r_{12}^p \left[\frac{\varepsilon_{12}}{(r_{12}^p - 1)} \right] \tag{13}$$

for p. This is simple for r constant (not necessarily 2 or integer), giving

$$p = \ln(\varepsilon_{23}/\varepsilon_{12})/\ln r. \tag{14}$$

(Note that Equation 14 differs slightly from Equation 11 but is equivalent to the order of the approximation involved.)

If r is not constant during the grid refinement, Equation 13 is transcendental in p. Usual solution techniques (e.g. Newton-Raphson) can apply, but one should allow for observed $p < 1$. This result can happen even for simple problems at least locally (de Vahl Davis 1983), and in some cases the observed $p < 0$ (unfortunately, behavior far away from asymptotic convergence can be nonmonotone). Also, $r \sim 2$ will be easier to solve than $r \sim 1 + \delta$, and $r \gg 2$ is probably not of much interest. For well-behaved synthetic cases, simple substitution iteration with a relaxation factor $\omega \sim 0.5$ works well. With $\rho =$ previous iterate for p, the iteration equation is

$$p = \omega\rho + (1 - \omega)\frac{\ln(\beta)}{\ln(r_{12})} \tag{15a}$$

$$\beta = \frac{(r_{12}^\rho - 1)}{(r_{23}^\rho - 1)} \frac{\varepsilon_{23}}{\varepsilon_{12}}. \tag{15b}$$

Note this form of the iteration gives the exact answer in one step for the case of $r = 2$ and $\omega = 0$.

Predicting the Required Grid Resolution

Once p is known with some confidence, one may predict the next level of grid refinement r^* necessary to achieve a target accuracy, expressed as a target error estimate E_1 or a target GCI_1 called GCI^*. With GCI_{23} being the value from Equation 2 for the previous two grids,

$$r^* = \sqrt[p]{GCI^*/GCI_{23}}. \tag{16}$$

This result, of course, depends only on the assumed definition of order of the discretization error, that is, only on $c = error/\Delta^p$ and not on the GCI theory itself.

Special Considerations for Turbulence Modeling

Special considerations are required for turbulence modeling and for other fields with multiple scales. Here, the code theoretical performance can be verified (within a tolerance) for a range of parameters but could fail in another range.

It is necessary to get the grid resolution into the asymptotic range in order to do grid-convergence testing. Virtually any grid is in the asymptotic range for a simple Laplace equation. For any boundary-layer calculation, it is clear that the initial (coarse) grid must get some points into the boundary layer. For turbulence modeling without wall functions (Wilcox 1993, Shirazi & Truman 1989) the grid must get some points into the wall layer. For turbulence modeling with wall functions, the grid should not get into the wall layer (Celik & Zhang 1995). In my interpretation, the wall functions should be viewed as an elaborate nonlinear boundary condition, and the grid-convergence exercise should be done from the edge of the wall layer out. Similarly, for large eddy simulations (LES), as used in aerodynamic turbulence research and in atmospheric and ocean modeling with sub-grid turbulence modeling, the grid convergence must not go to zero, or else the Reynolds stresses will be counted twice, once from the full Navier-Stokes terms and once modeled from the LES terms. Also, the presence of any switching functions, such as length determinations for the Baldwin-Lomax turbulence model (Wilcox 1993) can easily corrupt second-order convergence rates.

Finally, the grid-resolution requirements are much more demanding for turbulent boundary layers, just as laminar boundary layers are much more demanding than inviscid flows. For example, Claus & Vanka (1992) found that 2.4 million nodes ($256 \times 96 \times 96$) did not demonstrate grid independence of the computed velocity and turbulence fields of crossflow jets.

Special Considerations for Artificial Dissipation

Special considerations are required for any numerical treatment that effectively changes the governing PDEs as the grid is refined. Sub-grid turbulence modeling was mentioned earlier. The other important but troublesome numerical methods are those that treat shock-like problems with any kind of artificial dissipation that depends on the grid size. (*Shock-like problems* means either true shocks or simply regions of unresolvable large solution structure involving solution gradients, curvatures, and various combinations.) This includes both explicit artificial dissipation (in which an identifiable additional term is added to the governing PDEs, whether treated explicitly or implicitly in the solution algorithm) and implicit dissipation introduced locally via nonlinear flux limiters (FCT, TVD, ENO, PPM, and so on). If not addressed, these terms corrupt the convergence behavior. See Kuruvila & Anderson (1985) for discussion of difficulties and pitfalls of doing convergence studies with artificial dissipation

terms in the equations. The explicit dissipation approach can be treated correctly and rigorously, as in Blottner (1990) and Shirazi & Truman (1989).

The artificial viscosity of first-order upstream differencing (Roache 1972) requires no special consideration other than its slow convergence. However, hybrid methods in their various forms cause more problems. By adaptively changing the weighting given to first-order and second-order stencils, depending on the local cell Reynolds number Rc, these methods by definition are ultimately second-order accurate, but only when the resolution is so fine that the hybrid algorithm itself is inoperative, typically when $Rc < 2$. Over the practical range of resolution of interest, the order is ill-defined, but the empirically observed order is, as expected, somewhere vaguely between 1 and 2 (e.g. see Celik & Zhang 1995). Unfortunately, hybrid methods are commonly used in commercial codes, wherein the emphasis is often on code robustness at the sacrifice of accuracy. The same ill-defined behavior is observed for the class of Allen-Southwell methods (Roache 1972).

Error Estimation from Higher-Order Solutions on the Same Grid

Category B.1 methods estimate the accuracy of a base solution by comparison with "higher-order accuracy solution(s)." (Again, note the somewhat abusive but common terminology.) Richardson (1908) again scooped modern error-estimation papers by inventing Category B.1, error estimation from higher-order solutions on the same grid, noting that the difference between a second-order accurate solution and a fourth-order accurate solution is itself an ordered error estimator. The higher-order accuracy solution might be obtained via a new solution of higher-order discretizations using FDM (finite difference methods), FVM (finite volume methods), FEM (finite element methods), and so on, or by deferred corrections, compact differences (again direct or deferred corrections, and so on).

For the technique of error estimation from higher-order solutions on the same grid, much the same advantages and limitations apply as with the grid-convergence technique. It applies to all point values and functionals (e.g. lift, drag); x and t errors may be estimated independently or coupled; and the error estimate includes nonlinear coupling. These methods are not commonly used because they require additional code capability, unlike grid-convergence tests (Category A). On the other hand, these Category B methods do not require additional grid generation.

The development costs of the additional code capability may be minimized by noting several points. If we were intending to use the higher-order solution itself, many restrictions and requirements could apply, such as full iterative convergence, strict conservation, and so on. However, if our only use of the

higher-order solution is to estimate the error of the base solution, these considerations are not so important. The point is that error estimation of, for example, a second-order solution using fourth order methods is less demanding numerically than obtaining a fourth-order solution for direct use. Also, though not generally recognized, directional splitting works. (The following development has not been completed for cross-derivative terms, which in any case may require some careful formulation for higher-order stencils and should be verified rigorously.)

We use a notation similar to that above for multiple grid solutions: f denotes the exact solution, f_2 the second-order accurate solution, and f_4 the fourth-order accurate solution; C_{2x} are the coefficients of the Taylor's theorem expansion for the second-order solution in the x-direction, and so on; and R_{2x} are the remaining terms in the complete second-order expansion in x, and so on. Then, in two dimensions,

$$f = f_2 + C_{2x}\Delta x^2 + R_{2x} + C_{2y}\Delta y^2 + R_{2y} \tag{17}$$

and

$$f = f_4 + C_{4x}\Delta x^4 + R_{4x} + C_{4y}\Delta y^4 + R_{4y}. \tag{18}$$

Defining the error of the second-order solution E_2 to be

$$E_2 = f - f_2 \tag{19}$$

and substituting for f from Equation 18, we obtain

$$E_2 = f_4 - f_2 + 0(\Delta^4). \tag{20}$$

This is the basic (Richardson 1908) result that the difference between the second- and fourth-order solutions on the same grid is itself a fourth-order error estimate for the second-order solution (a result so obvious that it hardly deserves to be named a theorem). Likewise obvious, if we define the error of the fourth-order solution E_4 to be

$$E_4 = f - f_4 \tag{21}$$

and substituting for f from Equation 18, we obtain

$$E_4 = f_2 - f_4 + 0(\Delta^2) \tag{22}$$

for a second-order error estimator for the fourth-order solution.

The somewhat less obvious result is for the directional splitting of the higher-order solutions, which can be much easier to implement than a fully directional higher-order solution. Let f_{4x} denote the solution obtained with fourth-order

discretization in x and second-order in y, and f_{4y} denote the solution obtained with fourth-order discretization in y and second-order in x. The C_{2x} coefficients are unchanged from the previous definitions. Then

$$f = f_{4x} + C_{4x}\Delta x^4 + R_{4x} + C_{2y}\Delta y^2 + R_{2y} \tag{23}$$

and

$$f = f_{4y} + C_{4y}\Delta y^4 + R_{4y} + C_{2x}\Delta x^2 + R_{2x}. \tag{24}$$

We now estimate the difference between f_{4x} and f_2, using Equations 23 and 17, respectively (if we dropped the higher-order remainder terms at this point, we would prove the resulting estimator to be second-order accurate, but by retaining these terms presently, we show that the estimator is fourth-order accurate)

$$\begin{aligned}
f_{4x} - f_2 &= f - C_{4x}\Delta x^4 - R_{4x} - C_{2y}\Delta y^2 - R_{2y} \\
&\quad - \{f - C_{2x}\Delta x^2 - R_{2x} - C_{2y}\Delta y^2 - R_{2y}\} \\
&= C_{2x}\Delta x^2 + R_{2x} - C_{4x}\Delta x^4 - R_{4x}.
\end{aligned} \tag{25}$$

Similarly, we estimate the difference between f_{4y} and f_2 using Equations 24 and 17, respectively:

$$\begin{aligned}
f_{4y} - f_2 &= f - C_{4y}\Delta y^4 - R_{4y} - C_{2x}\Delta x^2 - R_{2x} \\
&\quad - \{f - C_{2x}\Delta x^2 - R_{2x} - C_{2y}\Delta y^2 - R_{2y}\} \\
&= C_{2y}\Delta y^2 + R_{2y} - C_{4y}\Delta y^4 - R_{4y}.
\end{aligned} \tag{26}$$

Adding Equations 25 and 26 gives

$$\begin{aligned}
f_{4x} + f_{4y} - 2f_2 &= C_{2x}\Delta x^2 + R_{2x} + C_{2y}\Delta y^2 + R_{2y} \\
&\quad - C_{4x}\Delta x^4 - R_{4x} - C_{4y}\Delta y^4 - R_{4y}.
\end{aligned} \tag{27}$$

In comparison with Equation 17, the first four terms on the right-hand side of Equation 27 are identically equal to $f - f_2$, without approximation. (That is, the remainder terms R_{2x} and R_{2y} are included.) The last four terms on the right-hand side of Equation 27 are $O(\Delta^4)$. With the definition given in Equation 19, we have the error of the second-order solution E_2 to be estimated as

$$E_2 = f_{4x} + f_{4y} - 2f_2 + O(\Delta^4). \tag{28}$$

Thus the estimate for the second-order solution can be obtained to fourth order by directionally split fourth-order solutions.

Error Estimation from Lower-Order Solutions on the Same Grid

Certainly a lower-order solution on the same grid could be used to estimate the uncertainty of the higher-order solution (Category B.2), in the same way that grid coarsening can be used rather than grid refinement, but I know of no such applications.

Residual Evaluation and Transport

In one type of Category C method (auxiliary PDE solutions on the same grid) and in Category D methods as well, *residuals* of the solution are evaluated. In the Category C methods, these residuals are used as a source term to a linearized error equation that mimics the governing equations. That is, the errors are advected, diffused, and so on. In the various flavors of this approach, much is made of the distinction between "exact error residual equations" and "approximate error residual equations." There is less than meets the eye in this distinction, because somewhere in the development, additional approximations are required. Indeed, there can be no exact evaluation of the error of the discrete point-wise solution by passing it through the continuum operator(s) of the governing PDEs, since the domain of the PDEs requires more than point-wise values. The philosophy of FEM, or the underlying conceptual model of FEM practitioners, is usually distinct from FVM and FDM in this regard, often leading to disconnects in communications. To most FDM and FVM practitioners, the numerical solution consists of point-wise, discrete values of the flow variables (as well as related solution functionals, such as drag coefficients and so on, obtained by numerical quadratures of the discrete values.) To FDM and FVM people, the term *residual* ordinarily refers to the algebraic residual, the leftovers from the numerical solutions of the discrete algebraic equations, that will result from incomplete iterative solution or from accumulated round-off errors in a direct matrix solver. Therefore, it makes no sense to FDM or FVM people to speak about estimating the discretization error of a PDE solution, which involves real-number system arithmetic, by evaluating the (algebraic) residual, which may be made arbitrarily small no matter how crude and inaccurate the discretization may be. (In fact, the cruder the discretization, the easier it is to reduce this algebraic residual.)

FEM practitioners have a different viewpoint. They invariably conceptualize the numerical solution as more than point-wise, since the selection and construction of the inter-element *basis functions* is an essential part of the FEM methodology. Thus, they conceptualize their numerical solution as a continuum solution and, thence, are led to consider the (differential) residual resulting from passing this solution function through the PDE operator. When an FEM author talks about residuals, he or she refers to what I distinguish here as differential residuals. (Incidentally, I believe that this difference in attitude and semantics can be traced to the historical accident that FDM practice grew up with iterative solvers, whereas FEM practice grew up with direct solvers.) Note the major distinction between these approaches: The second (differential) residual would be nonzero even if the computer used true real-number arithmetic, but

the first (algebraic) residual actually can be driven to machine zero in many (not necessarily all) algorithms/codes.

Once this useful semantic distinction is made, we can examine the differences in the approach. The practices are indeed different, but the difference is not as essential as one might be led to think. First of all, we note that most (moderate-order) FEM solutions, although they produce functions, cannot be passed through the PDE operator everywhere because they are not smooth at element boundaries. This limitation can be sidestepped by using the weak or integral form of the governing equations. Note, however, that conceptualizing the solution as a continuum function engenders some problems. For example, the solution may not be truly (locally and globally) conservative, that is, conserving of mass, momentum, and/or energy, even though the conservation form (Roache 1972) of the equations was used. Also, when these FEM solutions are graphed for presentation of results, and often when they are postprocessed for evaluation of solution functionals such as drag coefficients, the basis functions are not used. The quadratures can use the same algorithms used for FDM solutions or for any point-wise data. Conversely, on the FDM/FVM side, it is not necessary to conceptualize the solution as consisting of only the discrete (node or volume) points. Indeed, a simple methodology for deriving a finite difference method involves fitting polynomials through node point values (e.g. a parabolic fit leads to standard second-order centered difference stencils). Therefore, there is no reason why FEM approaches to error estimation (either for solution adaptive grids or for quantification of uncertainty) cannot be applied to FDM solutions.

All of this discussion is preamble to the most interesting aspect of residual evaluation. Regardless of the approach (FEM, FDM, FVM), if a (differential) residual is to be evaluated for error evaluation, *it must be based upon a functional form distinct from that used in the discretization that produced the solution.* If not, the differential residual evaluation will be practically zero, except for some algebraic garbage caused by incomplete iterative convergence or machine round-off. Although arguments may be made about which functional form to use for the (differential) residual evaluation, they all essentially work for error estimation and, therefore, the process (*a*) is somewhat arbitrary, and (*b*) applies to FDM and FVM as well as FEM.

Note that each conceptualization of the underlying (sub-cell) functional form of the solution leads to a discretization, but the discretization is not necessarily unique to the conceptualization. For example, the variable coefficient diffusion problem can be conceptualized as having smooth property variation or piecewise constant property variation. (In groundwater flow and transport problems,

the first conceptualization typically is held by geophysicists and engineers, and the second by geologists.) Nevertheless, for a single grid, both conceptualizations can lead to the identical discretization and therefore identical discrete solution. This suggests some tolerance in the construction of sub-cell solutions. To put it another way, just because your point-wise solution agrees with experiment does not necessarily prove the literalness of your underlying conceptualization.

Auxiliary PDE Solutions on the Same Grid

The procedure for error estimation in Category C (auxiliary PDE solutions on the same grid) then involves selection of an alternate functional form for the intercell (interelement) solution and evaluation of the (differential) residual. Again, the derivatives may not exist at element boundaries. Whether one selects a C^2 smooth function to begin with, or blends (interpolates) the differentials obtained at element (or cell) centers to produce smooth functions over the entire region, or puts off such questions to the quadrature rules, etc, is not very significant. The residuals can then be used as local error estimators (Category D) or transported in an error equation (Category C), which is indeed a significant distinction. The major point is that the residual evaluation must involve a discretization rule different from that used in obtaining the solution. In this way, these methods are after all not so distinct from Category B methods, which use higher- (or lower-) order stencils to evaluate the error. In Category C, a higher-order stencil is used to obtain a higher-order solution, but if the stencil were used only locally to obtain a local residual, it would be like Category D. The approach in which a rigorously iterated solution is replaced with a less completely converged solution is intermediate, and the globalness of the error estimate becomes dependent on the amount and type of iteration.

For examples of the error transport equation approach, see Ferziger (1993) and Van Straalen et al (1995). In all cases, the error transport equation will be some linearized version of the governing equations. It will be somewhat cheaper to solve than the original fluid dynamics equations (e.g. full Navier-Stokes) but, as expected, will be a less accurate error estimator near separations and so on. This approach is truly global but not as reliable as Category A or B methods.

Auxiliary Algebraic Evaluations on the Same Grid; Surrogate Estimators

Category D methods, involving algebraic evaluations on the same grid, are cheap, need no additional grid generation, and use no significant dynamic memory. However, the error evaluated usually has no direct relation to any error measure of engineering or scientific interest. Rather, the error measure is defined for mathematical convenience (as are the L_2 and L_∞ norms) or by physical intuition (as is kinetic energy conservation). These measures can be used as a surrogate for measures of engineering and scientific interest if and

only if a correlation is established via numerical experimentation over a suite of problems and range of problem parameters.

Nonconservation of Conservation Variables

Category D.1 (nonconservation of conservation variables) applies only to codes that do not use fully conservative algorithms for conservation variables. For example, mass is not identically conserved in most pre-1980 boundary-layer codes, in many FEM codes, as well as other CFD codes. The error estimate then involves numerical evaluation by quadrature for the erroneous loss or gain of mass in the computational domain. If the coded algorithm is consistent, this error will approach zero but only in the limit of $\Delta \to 0$ for a nontrivial problem. (Even for a nominally mass-conserving full Navier-Stokes code, the satisfaction of mass conservation will usually depend on the degree of strict iterative convergence achieved, but this gives no indication of the discretization error of the solution, which is our interest here.) Momentum, vorticity, and internal energy are other possibilities. Note that the evaluation of this error depends on the accuracy of the quadrature, which probably should be consistent with the algorithm for solving the PDEs (see also discussion of Category D.4 below), but this does not appear to be a strict requirement. Conservative codes and algorithms are generally preferred, but nonconservative codes are seen to offer a readily evaluated and understood (although still surrogate) error measure.

Nonconservation of Higher Moments

Category D.2 involves evaluation of conservation errors for higher-order moments. In typical turbulent (Reynolds-Averaged) Navier-Stokes codes, fully conservative discretization may be used for the primitive variables of mass, momentum, and internal energy. However, turbulent kinetic energy is typically not identically conserved. An evaluation of its global and local conservation by quadrature, including carefully evaluated boundary inflow terms and dissipation (Haworth et al 1993; Chang & Haworth 1995, 1996), then gives a surrogate indication of general discretization errors. In Chang & Haworth (1995), it is also used to reliably guide local grid refinement for solution-adaptive grid generation.

In an earlier section, I noted that, if there is any suspicion that the grid resolution is not in the asymptotic range, three grid solutions are necessary to verify (or determine) the rate of convergence and thereby estimate the error, when the exact solution is not known. The qualifier used, "when the exact solution is not known," may seem redundant when the context is a realistic problem, but there is an important and useful distinction to be made in regard to Category D.1 and D.2 methods that are based on conservation errors. Although the conservation imbalance may not be of direct engineering or scientific relevance, it

has the advantage that the exact solution of this quantity is known, namely zero. Thus, in this surrogate error measure, one requires only a single grid solution to calculate the error if p is known, and only two grid solutions to extract the observed convergence rate p (see Chang & Haworth 1996).

Zhu-Zienkiewicz-Type Estimators

Category D.3 methods (the Zhu-Zienkiewicz-type estimators) were developed and intended primarily to guide grid adaptation. See Zhu & Zienkiewicz (1990), Zienkiewicz & Zhu (1987, 1992), and other FEM methods that have a similar flavor, such as those of Strouboulis & Oden (1990), Oden et al (1993), Babuska et al (1994), and Ewing et al (1990). These methods allow the global energy norm to be well estimated and give good evaluation of local errors (and provide local estimate of stress accuracy, certainly important for structures problems). More relevant in the present context of the quantification of uncertainty, these estimators can be used as surrogate error estimators, and they are relatively cheap.

Although developed for FEM, the Zhu-Zienkiewicz approach is adaptable to FDM and FVM (D Pelletier, personal communication), using the concepts discussed above in the section "Residual Evaluation and Transport."

The Zhu-Zienkiewicz-type methods share the shortcoming of all surrogate estimators for fluid dynamics—other than guidance for grid adaptation, there is little inherent engineering or scientific interest in the error measure as defined. Therefore, unless the only interest is mathematics for its own sake, it is necessary to establish a correlation of the Zhu-Zienkiewicz estimator with an error measure of interest. This can be accomplished only by expensive numerical experimentation for a given class of problems. For example, a correlation based on numerical experiments for internal combustion engine modeling would be unlikely to provide any guidance for external aerodynamics, nor even for a large range of flow parameters for geometrically similar problems. However, when one is involved in extensive suites of calculations, one may build up such correlations by experience and arrive at a practical and very inexpensive error estimator.

This need for establishing a correlation in error measures is also true if one direct error measure (e.g. obtained by a grid-convergence study) is to be used as a surrogate for another. For example, in an airfoil calculation, does a 1% error estimate on C_L ensure a 1% (or 5%, and so on) error estimate for C_M? The correlation providing the (fuzzy) answer will be valid only for a restricted range of angle of attack, Reynolds number, Mach number, and airfoil class.

Convergence of Higher-Order Quadratures

Category D.4 methods are the simplest, involving convergence of higher-order quadratures (numerical integration of a known function). For example, quadrature for pressure and shear force on the surface leads to evaluation of a drag

coefficient C_D. For nontrivial problems (i.e. problems with significant solution structure), second-order quadrature will give a different answer than fourth-order accurate quadrature. As the discretization refines, these two quadratures will converge to each other. For coarse grids, the difference may be used as a surrogate error estimator. The same philosophy can be used point-wise. The simplest approach ultimately reduces to comparing the calculated point-wise value of some variable of interest with the value interpolated between the point's neighbors. Indeed, these values will converge as the solution converges and, therefore, can be (and have been) used as surrogate error estimates. However, the poverty of the concept is shown by the fact that the error estimate is obtained from a single solution without recourse to the governing PDEs. It is difficult to take seriously an error estimate for point-wise pressures from an incompressible Navier-Stokes code when the same error estimate would result from the same point-wise values in an $M = 20$ solution or in Darcy flow in a porous medium. In fact, since the governing PDEs are not invoked, these error estimators would be better named simply *resolution indicators*.

Time Accuracy Estimation

Estimation of the numerical error of the time discretization can be done in the same manner as the spatial errors, but simpler methods may also be used. In fact, it is relatively straightforward to estimate the temporal error as the calculation evolves, and to build a code with a solution-adaptive time step to control the temporal error to a predetermined level, even when only first-order time differencing is used.

Consider the following inexpensive temporal error estimator (Oden et al 1993; Roache 1993c, d) for a fully implicit (backward) time differencing method. We write the backward time method for a general equation system in terms of an operator L (not necessarily linear) as

$$\left(f_{i,j}^{n+1} - f_{i,j}^n\right)/\Delta t = L^{n+1}. \tag{29}$$

Generally, the solution will involve an expensive matrix solution for all f^{n+1}. The time-error estimator uses the difference between backward and forward time integration, implemented as an extrapolation. The method is very cheap to implement because it does not require another implicit matrix solution, nor even another explicit stencil evaluation. It includes the effects of time-dependent boundary conditions and source terms.

In the current time-step, Equation 29 is advancing the solution for f from time level n to time level $(n + 1)$ with increment Δt using fully implicit (backward) time differencing, so that the operator L is being evaluated at $(n + 1)$. In the previous time-step, the solution was advanced from $(n - 1)$ relative to the

current indexing to n with increment Δt_{OLD} and the L evaluated at n. We could explicitly evaluate L^n and make a separate, parallel estimate of the values f^{n+1} with an explicit step, as in

$$\left(F_{i,j}^{n+1} - f_{i,j}^n\right)/\Delta t = L^n \tag{30}$$

where F^{n+1} signifies the new value of f predicted by the explicit algorithm. (Since the left-hand side involves values of f only at the location (i, j) but not at the neighboring locations $(i \pm 1, j)$ and $(i, j \pm 1)$ the equation is explicit, that is, it does not involve a matrix solution of f at all values of i and j simultaneously.) The difference between the new value of f^{n+1} predicted by the implicit algorithm and F^{n+1} predicted by the explicit algorithm, both of which predictions are $O(\Delta t)$ accurate, is itself an error estimator of accuracy $O(\Delta t^2)$ for the time discretization error for that time step. Explicit evaluation of L^n would not be expensive compared with the computer time necessary for the implicit matrix solution, but it does involve coding storage penalties and complexities (storage of old values of boundary conditions, source terms, and so on). An economical and elegant approach is to recognize that the L from the previous implicit step is identical to the L for the present explicit step. Thus, the L for the present explicit step can be evaluated from the knowledge of the previous change in f, requiring only the temporary storage of previous solution arrays. The explicit solution for f at $(n + 1)$, F^{n+1}, is then obtained by simple linear extrapolation of previous solutions. For constant Δt,

$$F_{i,j}^{n+1} = 2 f_{i,j}^n - f_{i,j}^{n-1}. \tag{31}$$

For the more general case of variable Δt,

$$F_{i,j}^{n+1} = f_{i,j}^n + \frac{\Delta t}{\Delta t_{OLD}}\left(f_{i,j}^n - f_{i,j}^{n-1}\right). \tag{32}$$

The point-wise difference $F^{n+1} - f^{n+1}$ is then a point-wise temporal error estimator. Usually, a user would be interested in the maximum over the spatial domain of the percentage error, so that one would evaluate the error estimator E_{ET} for the single time-step as

$$E_{ET} = 100 \cdot \max_{i,j}\left\{abs\left(F_{i,j}^{n+1} - f_{i,j}^{n+1}\right)\right\}/f_{RANGE}. \tag{33}$$

E_{ET} is thus calculated as the percent maximum deviation (or L^∞ norm) of the absolute value of the difference between the new f values f^{n+1}, predicted by the fully implicit algorithm, and F^{n+1}, predicted by the explicit algorithm, normalized by f_{RANGE}, which is the total range of f^{n+1}. (In practice, ghost-point evaluations with Dirichlet boundary conditions tend to exaggerate the error, so error estimates and f_{RANGE} should be calculated only over interior points.)

Note that the explicit calculation is used only as an error estimator within a time step of an implicit method, not as the solution algorithm. That is, its effects do not accumulate. Thus, stability limitations and/or conservation issues of explicit time-stepping as a solution algorithm are irrelevant.

Although the extrapolation procedure is equivalent to explicit time-stepping, the extrapolation cannot be started until there are two time levels. Also, in the event that the initial conditions are set arbitrarily by the analyst (without setting initial conditions as a steady-state solution), the initial conditions are likely to be incompatible with the boundary conditions applied at the first time-step. This means that the change in boundary values during the first time-step is approximately fixed, that is, it does not depend much on the time resolution. Consequently, the first time-step would not provide a meaningful estimate of $\partial f/\partial t$, and the error estimator would be invalid.

E_{ET} can be used readily (internal to the code) as the basis for a solution-adaptive time-stepping algorithm, adjusting Δt so that the error estimate E_{ET} is acceptable, either by recalculating the previous time-step (preferred) or by simply adjusting the next time-step.

Nonlinear Dynamics Solutions

Yee et al (1991) and Lafon & Yee (1992) have given a large number of examples of nonlinear (chaotic) dynamics solutions in which spurious (grossly erroneous) solutions for steady-state problems with strongly nonlinear source terms are obtained for stable implicit algorithms applied beyond the linear stability limit. Their point is the danger of being misled by such calculations, and the need for further research to guarantee *a priori* accuracy. I believe the real lesson of the examples is simply the need for *a posteriori* error estimation. In all but one of the examples, the most cursory time-step convergence study would quickly reveal the inadequacy of the temporal resolution, with no reasonable chance of being misled into a nonphysical solution. The one remaining example (involving a fourth-order Runge-Kutta integration) could be misleading with a cursory study but would be revealed in a thorough systematic study. However, it is worth noting and is perhaps nonintuitive that even steady-state solutions may require time-step resolution (or equivalently, relaxation parameter resolution) when significant nonlinearities are present.

Cafe Curves

Different fields of endeavor have different requirements for the variables and functionals of interest. To generalize too much, engineering (mechanical, aerospace, chemical) often requires functionals of the solution such as lift coefficient, heat transfer rates, mixing rate, and so on. It is much less common that practical interest is in accuracy of the entire field calculation; one example that

comes to mind is aero-optical propagation (Truman & Lee 1990). However, meteorological and ocean modelers are usually interested in accuracy of all the field variables. In these cases, it is meaningful to emphasize global accuracy of the entire computational field, and difficulty arises in attempting to characterize global accuracy with just one or a few numbers.

Often L_2 or L_∞ error norms are used as global error indicators. However, questions arise about whether these error norms are always reliable, and more significantly how they can be interpreted in any general sense. For example, in a tidal convergence study (Luettich & Westerink 1995, Westerink et al 1994), L_∞ norms are poor indicators of error because shifting amphidromes, associated with very high gradient in response amplitude and phase, cause extremely high localized errors that do not converge smoothly or at the same rate as an L_2 norm. Furthermore, there is the general problem of interpreting how representative an L_2 or L_∞ norm is to the response as a whole, or to aspects of the solution of particular interest.

Luettich & Westerink (1995) introduced the concept of cumulative area fraction error (CAFE) curves to present domain errors in a more complete and meaningful way. These curves plot the fraction of the total domain that exceeds a particular error level (y-axis) against that error level (x-axis). Under- and over-prediction are indicated separately. Assuming that under- and over-prediction are distributed approximately evenly throughout the domain, median under- and over-prediction errors correspond approximately to a cumulative area fraction of 0.25. For an example of the cumulative-area-fraction error curves, see the grid intercomparisons for tidal flow computations reported by Westerink et al (1994); see also Westerink & Roache (1995).

The Myth of the Converged Solution

CFD practitioners, especially industrial engineers using general-purpose commercial CFD codes, would like to be able to obtain, at least in principle, a converged solution (converged both iteratively and in time-space discretization) and then be able to confidently postprocess that solution for any engineering measure of interest. Unfortunately, this is not possible in any general sense. Again, the problem is the *a priori* unknown correlation between possible error measures of valid engineering and scientific interest. Of course, expert judgment (based on experience with a nearby class of problems) can be valid, but the problem is not soluble in any general sense.

For example, in our experience with dynamic stall and oscillating airfoil and wing calculations (Salari & Roache 1990, Salari et al 1994), lift was well converged at the finest discretization achieved ($141 \times 55 \times 55$ in 3-D, and 461×71 in 2-D), drag was more problematical, and moment definitely was not converged. If one had performed the grid-convergence study (or other

error-estimation exercise) examining only the lift, one could only infer from engineering judgment (experience on related problems) whether or not resolution was adequate for drag and moment. If a new error measure became of interest, say position of transition, or location of trailing-edge separation, or second harmonic component of the unsteady pressure coefficient (of interest for helicopter dynamic stall analysis; McCroskey 1981), the convergence study would have to be repeated for this error measure, strictly speaking. Calculations of the branching between symmetric and nonsymmetric flow patterns require special attention to grid resolution (see Sengupta et al 1995). Rosenfeld (1994) used grid resolution studies of bluff-body wakes with resolution up to 513×513 points, and examined convergence in both the physical and Fourier domains; he showed that phase velocities of vortices converge much slower than amplitudes, so that coarser grids may be acceptable for prediction of force coefficients.

Certainly, it is always possible to devise some error measure that is exquisitely sensitive to discretization error, for example, some high-order statistical correlations in turbulent flow calculations, so that this measure is far from converged even when other, more benign measures are well converged. A practical example is a boundary-layer stability calculation. For most engineering applications, skin friction and wall-heat transfer are of principal interest, and these are sensitive to the first normal derivative of velocity at the wall. However, it is known from stability theory (Lin 1967) that laminar-boundary layer stability is dependent on the diffusion of vorticity across the critical layer (the y-position at which the mean flow speed equals the disturbance wave speed). The principal component of boundary-layer vorticity magnitude is $\partial u/\partial y$, which means that the appropriate measure of accuracy is of the term $\partial^3 u/\partial y^3$, which may be expected to be more difficult to converge than simply the velocity component u. Conversely, accurate convergence of pressure may be unnecessary.

Thus, the concept of a converged solution, ascertained to be so independent of the intended error measure, is a myth.

Error Estimation for Grid Adaptation versus Quantification of Uncertainty

Local error estimators are used for solution-adaptive grid-generation algorithms, and they are usually successful for this purpose (e.g. see Schonauer et al 1981, Oden et al 1993). However, almost anything intuitive is successful for adaptation purposes, e.g. minimizing solution curvature or adapting to solution gradients (even though solution gradients *per se* cause no error in most discretization schemes). In 1-D problems (e.g. Salari & Steinberg 1994) or quasi-1-D problems (Dwyer et al 1980), the gains in computational efficiency from solution adaptivity of the r-type (redistribution) in structured grids are very impressive. In real multidimensional problems, gains are usually modest

(Roache et al 1984, Hall & Zingg 1995) but are somewhat more significant for unstructured grid adaptation (e.g. see Morgan et al 1991, Lohner 1989, Hetu & Pelletier 1992, Pelletier & Ignat 1995).

But this task of grid adaptation has little connection to the quantification of uncertainty for a final calculation with a useful error measure. So the success of these local error estimators in guiding grid adaptation must not be taken as demonstration of their efficacy for the quantification of uncertainty. Estimates of local errors usually are not what we want for verification of calculations; we need global errors. By *global* I do not mean just a global summing up of local values (as is often used in the FEM literature) but an evaluation that includes nonlocal effects, that is, one taking into account the fact that errors are advected, diffused, and so on. For error estimation of useful scientific or engineering measures, local estimates are suspect, and their validity as surrogate estimators for measures of interest must be established anew for each family of nearby problems.

Internet Archive

An Internet archive of a bibliography for CFD Verification and Validation is maintained by R Barron of the University of Windsor, Ontario, Canada, and can be accessed at the following address: http://www.lpac.ac.uk/SEL-HPC/Articles/ValCFD.html

Final Remarks

The various approaches to error estimation and quantification of uncertainty in CFD have their relative merits, involving algebraic simplicity of local estimators vs the complexity of additional PDE formulations or higher-order accurate solutions, single-grid vs multiple-grid generations, postprocessing vs CFD code modifications, and so on. Systematic grid-convergence studies are the most common, most straightforward and arguably constitute the most reliable technique for the quantification of numerical uncertainty. This approach requires no code modifications nor algorithm developments but does require multiple grid generations. Significantly, it is not necessary to "double the grid"; noninteger grid refinement and grid coarsening are economical alternatives.

However, regardless of each approach's relative merits and of the possibility for future improvements, it is clear that methods are available now to convincingly assess the numerical uncertainty of CFD calculations. [It is certainly not necessary, as claimed in a recent SIAM article by Johnson et al (1995), to analytically and *a priori* solve the hydrodynamics stability problem in order to estimate CFD numerical errors!] Journal editors and reviewers can insist on reasonably thorough assessment without placing impractical demands on authors, to the improvement of journal quality.

Visit the *Annual Reviews home page* at
http://www.annurev.org.

Literature Cited

Aeschliman DP, Oberkampf WL, Blottner FG. 1995. A proposed methodology for computational fluid dynamics code verification calibration and validation. *Proc. Int. Congr. Instrumentation Aerospace Simulation Facilities.* Dayton, Ohio: Wright-Patterson Air Force Base

AIAA. 1994. Policy statement on numerical accuracy and experimental uncertainty. *AIAA J.* 32(1):3 (editorial)

ASME Editorial Board. 1994. *Journal of Heat Transfer* editorial policy statement on numerical accuracy. *ASME J. Heat Transfer* 116(11):797–98

Babuska I, Strouboulis T, Upadhyay CS. 1994. A model study of the quality of a posteriori error estimators for linear elliptic problems. Error estimation in the interior of patchwise uniform grid of triangles. *Comp. Meth. Appl. Mech. Engineer.* 114:307–78

Blottner FG. 1990. Accurate Navier-Stokes results for the hypersonic flow over a spherical nosetip. *AIAA J. Spacecraft Rockets* 27(2):113–22

Boehm BW. 1981. *Software Engineering Economics.* New York: Prentice-Hall

Celik I, Chen CJ, Roache PJ, Scheurer G, eds. 1993. *Symposium on Quantification of Uncertainty in Computational Fluid Dynamics.* FED-ASME, Vol. 158. New York: Am. Soc. Mech. Eng.

Celik I, Freitas CJ, eds. 1990. *Benchmark Test Cases for Computational Fluid Dynamics.* FED-ASME, Vol. 93. New York: Am. Soc. Mech. Eng.

Celik I, Zhang WM. 1995. Calculation of numerical uncertainty using Richardson extrapolation: application to some simple turbulent flow calculations. *ASME J. Fluids Engineer.* 117:439–45

Chang S, Haworth DC. 1995. Kinetic-energy-balance based solution-adaptive mesh refinement. See Johnson & Hughes 1995, pp. 7–12

Chang S, Haworth DC. 1996. Adaptive grid refinement using cell-level and global imbalances. *Int. J. Num. Meth. Fluids.* In press

Claus RW, Vanka SP. 1992. Multigrid calculations of a jet in crossflow. *J. Propulsion Power* 8:185–93

Cosner RR. 1995. CFD Validation requirements for technology. *AIAA 26th Fluid Dynamics Conf.* San Diego, CA

de Vahl Davis G. 1983. Natural convection of air in a square cavity: a bench mark numerical solution. *Int. J. Num. Meth. Fluids* 3(3):249–64

Dietrich DE, Roache PJ, Marietta MG. 1990. Convergence studies with the Sandia Ocean Modeling System. *Int. J. Num. Meth. Fluids* 11:127–50

Dwyer HA, Kee R, Sanders B. 1980. Adaptive grid method for problems in fluid mechanics and heat transfer. *AIAA J.* 18:1205–12

Ethier CR, Steinmen DA. 1994. Exact fully 3D Navier-Stokes solutions for benchmarking. *Int. J. Num. Meth. Fluids* 19:369–75

Ewing RE. 1990. A posteriori error estimation. *Comp. Meth. Appl. Mech. Engineer.* 82:59–72

Ferziger JH. 1993. Estimation and reduction of numerical error. See Celik et al 1993, p. 18

Freitas CJ. 1993a. Editorial policy statement on the control of numerical accuracy. *ASME J. Fluids Engineer.* 115(3):339–40

Freitas CJ, ed. 1993b. *The CFD Triathlon: Three Laminar Flow Simulations by Commercial CFD Codes.* FED-ASME, Vol. 160. New York: Am. Soc. Mech. Eng.

Freitas CJ. 1994. Response: To the Comments by Drs. W. Shyy and M. Sindir. *ASME J. Fluids Engineer.* 116(2):198

Freitas CJ. 1995a. Perspective: Selected benchmarks from commercial CFD codes. *ASME J. Fluids Engineer.* 117(2):208–18

Freitas CJ. 1995b. The end of justification, the beginning of implementation. *ASME J. Fluids Engineer.* 117(1):9

Gresho PM, Taylor C. 1994. *Int. J. Num. Meth. Fluids* 19:iii (editorial)

Hall DJ, Zingg DW. 1995. Viscous Airfoil Com-

putations Using Adaptive Grid Redistribution. *AIAA Journal* 33 (7):1205-10

Haworth DC, El Tahry SH, Huebler MS. 1993. A global approach to error estimation and physical diagnostics in multidimensional computational fluid dynamics. *Int. J. Num. Meth. Fluids* 17:75–97

Hetu JF, Pelletier DH. 1992. Adaptive remeshing for viscous incompressible flows. *AIAA J.* 30:1986–92

Horgan J. 1995. From complexity to perplexity. *Sci. Am.* 272(6):104–9

Johnson C, Rannacher R, Boman M. 1995. Numerics and hydrodynamic stability: toward error control in computational fluid dynamics. *SIAM J. Num. Anal.* 32:1058–79

Johnson RW, Hughes ED, eds. 1995. *Quantification of Uncertainty in Computational Fluids Dynamics–1995. FED-ASME*, Vol. 213. New York: Am. Soc. Mech. Eng.

Karniadakis GE. 1995. Toward a numerical error bar in CFD. *ASME J. Fluids Engineer.* 117(1):7–9

Knupp P, Steinburg S. 1994. *Fundamentals of Grid Generation.* Boca Raton: CRC

Konikow LF, Bredehoeft JD. 1992. Groundwater models cannot be validated. *Adv. Water Res.* 15:75–83

Kuruvila G, Anderson J. 1985. A study of the effects of numerical dissipation on the calculations of supersonic separated flows. *AIAA Paper 85-0301*

Lafon A, Yee HC. 1992. On the numerical treatment of nonlinear source terms in reaction-convection equations. *AIAA Paper 92-0419*

Leijnse A, Hassanizadeh SM. 1994. Model definition and model validation. *Adv. Water Res.* 17:197–200

Leonard BP. 1995. Comments on the policy statement on numerical accuracy. *ASME J. Fluids Engineer.* 117(1):5–6

Lin CC. 1967. *The Theory of Hydrodynamic Stability.* London: Cambridge Univ. Press

Lohner R. 1989. Adaptive remeshing for transient problems. *Comp. Meth. Appl. Mech. Engineer.* 61:323–38

Lotz RD, Thompson BE, Konings CA, Davoudzadeh F. 1995. Numerical uncertainties in transonic flow calculations for airfoils with blunt trailing edges. *Int. J. Num. Meth. Fluids.* In press

Luettich RA, Westerink JJ. 1995. Continental shelf scale convergence studies with a barotropic tidal model. In *Quantitative Skill Assessment for Coastal Ocean Models*, ed. DR Lynch, AM Davies. Washington: Am. Geophys. Union. In press

McCroskey WJ. 1981. *The Phenomenon of Dynamic Stall. NASA TM-81264*

Mehta U. 1989. Computational requirements for hypersonic flight performance estimates. *AIAA Paper 89-1670*

Mehta U. 1991. Some aspects of uncertainty in computational fluid dynamics results. *ASME J. Fluids Engineer.* 113:538–43

Mehta U. 1995. Guide to credible computational fluid dynamics simulations. *AIAA Paper 95-2225*

Melnik RE, Siclari MJ, Marconi F, Barger T, Verhoff A. 1995. An overview of a recent industry effort at CFD code certification. *AIAA Paper 95-2229*

Morgan K, Peraire J, Peiro J, Hassan O, Probert EJ. 1991. Unstructured Grid Methods for Compressible Flows. In *Computational Mechanics*, ed. YK Cheung, JHW Lee, AYY Leung, 2:1489–97. Rotterdam: Balkema

Nguyen TV, Maclaine-Cross IL. 1988. Incremental pressure drop number in parallel-plate heat exchangers. *ASME J. Fluids Eng.* 110:93–96

Oberkampf WL. 1994. A proposed framework for computational fluid dynamics code calibration/validation. *AIAA Paper 94-2540*

Oberkampf WL, Blottner FG, Aeschliman DP. 1995. Methodology for Computational Fluid Dynamics Code Verification/Validation. *AIAA Paper 95-2226*

Oden JT, Kennon SR, Tworzydlo WW, Bass JM, Berry C. 1993. Progress on adaptive h-p finite element methods for the incompressible Navier-Stokes equations. *Comput. Mech.* 11:421–32

Oldenburg CM, Pruess K. 1995. Dispersive transport dynamics in a strongly coupled groundwater-brine flow system. *Water Resources Res.* 31(2):289–302

Oreskes N, Shrader-Frechette K, Belitz K. 1994. Verification validation and confirmation of numerical models in the earth sciences. *Science* 263:641–46

Organization for Economic Cooperation and Development. 1988. *The International HYDROCOIN Project Groundwater Hydrology Modelling Strategies for Performance Assessment of Nuclear Waste Disposal Level 1: Code Verification. OECD Report 71617.* Paris: OEDC

Organization for Economic Cooperation and Development. 1990. *PSACOIN Level 1A Intercomparison.* Paris: OECD

Patankar SV. 1980. *Numerical Heat Transfer and Fluid Flow.* New York: McGraw-Hill

Pelletier D, Ignat L. 1995. On the accuracy of the grid convergence index and the Zhu-Zienkiewicz error estimator. See Johnson & Hughes 1995, pp. 31–36

Polya G. 1957. *How to Solve It; A New Aspect of Mathematical Method.* Princeton, NJ: Princeton Univ. Press

Putre M. 1994. Verifying analytical models. *ASME Mech. Engineer.* 116(11):70–74

Richardson LF. 1908. The approximate arithmetical solution by finite differences of physical problems involving differential equations with an application to the stresses in a masonry dam. *Trans. R. Soc. London* Ser. A, 210:307–57

Richardson LF. 1927. The deferred approach to the limit. *Trans. R. Soc. London* Ser. A 226:229–361

Roache PJ. 1972. *Computational Fluid Dynamics.* Albuquerque, NM: Hermosa. Rev. 1976

Roache PJ. 1982. Semidirect/Marching Methods and Elliptic Grid Generation. In *Proceedings Symposium on the Num. Generation of Curvilinear Coordinate Systems and Use in the Numerical Solution of Partial Differential Equations*, ed. JF Thompson, pp. 727–37. Amsterdam: North-Holland

Roache PJ. 1990. Need for control of numerical accuracy. *J. Spacecraft Rockets* 27(2):98–102

Roache PJ. 1993a. A method for uniform reporting of grid refinement studies. See Celik et al 1993, pp. 109–20

Roache PJ. 1993b. A method for uniform reporting of grid refinement studies. *Proc. 11th AIAA Comput. Fluid Dynam. Conf.*, Orlando, FL, 2:1057–58

Roache PJ. 1993c. The SECO code algorithms for groundwater flow and transport. In *Finite elements in fluids: new trends and applications. Proc. VIII Int. Conf. Finite Elements Fluids*, ed. K Morgan, E Onate, J Periaux, J Peraire, OC Zienkiewicz, 2:939–48. Barcelona, Spain: Univ. Politecnica Cataluna

Roache PJ. 1993d. The SECO suite of codes for site performance assessment. In *Proceedings 1993 Intl. High-Level Radioactive Waste Management Conf.* Las Vegas, NV

Roache PJ. 1994a. Perspective: a method for uniform reporting of grid refinement studies. *ASME J. Fluids Engineer.* 116(3):405–13

Roache PJ. 1994b. Response: To the Comments by Drs. W. Shyy and M. Sindir. *ASME J. Fluids Engineer.* 116(2):198–99

Roache PJ. 1995. Verification of codes and calculations. *AIAA Paper 95-2224;AIAI J.* In press

Roache PJ, Knupp PM. 1993. Completed Richardson extrapolation. *Comm. Appl. Num. Methods* 9:365–74

Roache PJ, Salari K. 1990. Weakly compressible Navier-Stokes Solutions with an implicit approximate factorization code. *AIAA Paper 90-0235*

Roache PJ, Ghia K, White F. 1986. Editorial policy statement on the control of numerical accuracy. *ASME J. Fluids Engineer.* 108(1):2

Roache PJ, Knupp PM, Steinberg S, Blaine RL. 1990. Experience with benchmark test cases for groundwater flow. See Celik & Freitas 1990, pp. 49–56

Roache PJ, Steinberg S, Moeny WM. 1984. Interactive electric field calculations for lasers. *AIAA Paper 84-1655*

Rosenfeld M. 1994. Grid refinement test of time-periodic flows over bluff bodies. *Comp. Fluids* 23(5):693–709

Salari K, Roache PJ. 1990. The influence of sweep on dynamic stall produced by a rapidly pitching wing. *AIAA Paper 90-0581*

Salari K, Steinberg S. 1994. Flux-corrected transport in a moving grid. *J. Comput. Physics* 111(1):24–32

Salari K, Blaine RL, Economy K, Roache PJ. 1995. Grid resolution studies of radionuclide transport in fractured porous media. See Johnson & Hughes 1995, pp. 25–30

Salari K, Roache PJ, Wilcox DC. 1994. *Numerical Simulation of Dynamic Stall Using the Two-Equation K-ω Turbulence Model*, ed. K Ghia, M Hafez. ASME-FED, 196:1–7. New York: Am. Soc. Mech. Eng.

Schonauer W, Raith K, Glotz G. 1981. The principle of difference of difference quotients as a key to the self adaptive solution of nonlinear partial differential equations. *Comp. Meth. Appl. Mech. Engineer.* 28(3):327–59

Sengupta R, Orkwis PD, Davis SM. 1995. Unsteady solutions transients and local grid resolution requirements for asymmetric vortex formation. *AIAA Paper 95-0364*

Shih TM, Tan CH, Hwang BC. 1989. Effects of grid staggering on numerical schemes. *Third Int. J. Num. Meth. Fluids* 9(2):193–212

Shirazi SA, Truman CR. 1989. Evaluation of algebraic turbulence models for PNS predictions of supersonic flow past a sphere-cone. *AIAA J.* 27(5):560–68

Shyy W, Sindir M. 1994. Comments on "Policy statement on the control of numerical accuracy." *ASME J. Fluids Engineer.* 116(2):196–97

Steinberg S, Roache PJ. 1985. Symbolic manipulation and computational fluid dynamics. *J. Comput. Physics* 57(2):251–84

Stephens AB, Shubin GR. 1981. Multiple solutions and bifurcation of finite difference approximations to some steady problems in fluid dynamics. *SIAM J. Sci. Stat. Comput.* 2(4):404–15

Strouboulis T, Oden JT. 1990. A posteriori error estimation of finite element approximations in fluid mechanics. *Comp. Meth. Appl. Mech. Engineer.* 78:201–42

Sudicky EA, Frind EO. 1982. Contaminant transport in fractured porous media: analytical solutions for a system of parallel fractures. *Water Resources Res.* 18(6):1634–42

Truman CR, Lee MJ. 1990. Effects of organized turbulence structures on the phase distortion in a coherent optical beam propagating through a turbulent shear flow. *Physics*

Fluids A 2(5):851–57

Vanka P. 1994. Response: To the Comments by Drs. W. Shyy and M. Sindir. *ASME J. Fluids Engineer.* 116(2):197–98

Vanka P. 1995. Comment on "Toward a numerical error bar in CFD." *ASME J. Fluids Engineer.* 117(1):9–10

Van Straalen BP, Simpson RB, Stubley GD. 1995. A posteriori error estimation for finite volume simulations of fluid flow transport. In *Proc. Third Annu. Conf. CFD Soc. Canada*, Vol. 1, ed. PA Thibault, DM Bergeron. Banff, Alberta

Westerink JJ, Luettich RA, Muccino JC. 1994. Modeling tides in the western North Atlantic using unstructured graded grids. *Tellus* 46A:187–99

Westerink JJ, Roache PJ. 1995. Issues in convergence studies in geophysical flow computations. *Soc. Industr. Appl. Math.* In press

Wilcox DC. 1993. *Turbulence Modeling for CFD*. La Cañada, CA: DCW Industries

WIPP PA Dept. 1992. *Annual Performance Assessment for the Waste Isolation Pilot Plant Dec. 1992 Vol. 1 Preliminary Comparison with 40 CFR 191 Subpart B Vol. 2: Technical Basis.* Albuquerque, NM: Performance Assess. Dept., Sandia Natl. Lab.

Yee HC, Sweby PK, Griffiths DF. 1991. Dynamical approach study of spurious steady-state numerical solutions of nonlinear differential equations. I. The dynamics of time discretization and its implications for algorithm development in computational fluid dynamics. *J. Comput. Physics* 97:249–310

Zhu JZ, Zienkiewicz OC. 1990. Superconvergence recovery technique and a posteriori error estimates. *Int. J. Num. Meth. Engineer.* 30:1321–39

Zienkiewicz OC, Zhu JZ. 1987. A simple error estimator and adaptive procedure for practical engineering analysis. *Int. J. Num. Meth. Engineer.* 24:337–57

Zienkiewicz OC, Zhu JZ. 1992. The superconvergent patch recovery and a posteriori error estimates. I: The recovery technique. *Int. J. Num. Meth. Engineer.* 33:1331–64

Zingg DW. 1993. Grid studies for thin-layer Navier-Stokes computations of airfoil flowfield. *AIAA J.* 30(10):2561–64