INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
Int. J. Numer. Meth. Fluids 2006; **50**:1335–1355
Published online 15 November 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/fld.1090

Quantitative V&V of CFD simulations and certification of CFD codes

Fred Stern*,†, Robert Wilson‡ and Jun Shao§

IIHR-Hydroscience and Engineering, The University of Iowa, 300 South Riverside Drive, Iowa City, IA 52242, U.S.A.

SUMMARY

Definitions and equations are provided for the quantitative assessment of numerical (verification) and modelling (validation) errors and uncertainties for CFD simulations and of intervals of certification for CFD codes. Verification, validation, and certification methodology and procedures are described. Examples of application of quantitative certification of RANS codes are presented for ship hydrodynamics. Opportunities and challenges for achieving consensus and standard V&V and certification methodology and procedures are discussed. Copyright © 2005 John Wiley & Sons, Ltd.

KEY WORDS: verification; validation; certification; CFD

INTRODUCTION

In spite of the ever-increasing need and importance for standards for computational fluid dynamics (CFD) uncertainty analysis/accuracy estimation, there are currently many viewpoints covering all aspects from basic concepts and definitions to detailed methodology and procedures. A similar situation existed for experimental fluid dynamics (EFD) uncertainty analysis, ca. 1960 for which currently standards are widely accepted and available, although widespread use is still lacking.

Pioneering work was done by Roache [1] who proposed the grid convergence index (GCI) for estimating uncertainty due to grid and time step errors based on Richardson extrapolation (RE) using multiple solutions on systematically refined grids, thereby, providing a quantitative metric for verification. Roache [2] expanded on this work through overall discussion of

*Correspondence to: Fred Stern, IIHR-Hydroscience and Engineering, The University of Iowa, 300 South Riverside Drive, Iowa City, IA 52242, U.S.A.
†E-mail: frederick-stern@uiowa.edu
‡E-mail: robert-wilson@utc.edu
§E-mail: jun-shao@uiowa.edu

Contract/grant sponsor: Office of Naval Research; contract/grant number: N00014-01-1-0073

Received 15 March 2005
Revised 11 August 2005
Accepted 14 August 2005

verification and validation (V&V), including: use of EFD definitions for errors and uncertainties; phrases (e.g. verification deals with equations solved correctly and validation with correct equations and verification deals with mathematics and validation with physics) and activities (e.g. method of manufactured solutions and benchmark solutions for verification and use of experimental data for validation) defining V&V; discussion for V&V of both codes and solutions; many case studies demonstrating use of GCI; single grid error estimation methods; and broader issues such as code quality assurance and certification. Such definitions for V&V, however, in the authors' viewpoint are inadequate. Quantitative metrics are needed for both verification and validation, and methodology is needed for combining errors and uncertainties.

The AIAA Committee on Standards for CFD [3] and Guide for V&V of CFD Simulations [4] uses definitions from information theory for errors and uncertainties with emphasis on measurement of accuracy as opposed to estimation of errors and uncertainties; follows Roache's phrases and expands considerably on his activities along with broad statements in defining V&V; discusses mostly code, but also solution V&V; and additionally discusses policy statements on experimental and numerical accuracy. Code verification activities measure accuracy in relation to benchmark analytical and ordinary and partial differential equation solutions for simplified problems along with software quality assurance: identify, quantify, and reduce errors in the computational model and its numerical solution. Model validation activities measure accuracy in relation to experimental data with emphasis on validation tiers based on unit problems, benchmark cases, subsystem cases, and complete systems: identify and quantify error and uncertainty in the conceptual and computational models, quantify the numerical error in the computational solution, estimate the experimental uncertainty, and compare the computational and experimental results. Solution verification largely follows Roache in using GCI along with consistency and iterative convergence checks. Rigorous implementation is impressive [5, 6]; nonetheless, these definitions are subject to same criticisms mentioned earlier. Another problem with such definitions is lack of an overall mathematical framework for V&V, which is considered essential in the author's viewpoint similarly as it is an essential and integral part of EFD uncertainty analysis.

The literature also includes editorial policy statements [7], additional guidelines [8], and numerous case studies, which mostly focus on verification procedures for 2D problems (e.g. volume 36 of the AIAA Journal and volume 124 of the ASME Journal of Fluids Engineering). In general, this literature follows approaches similar to that described above.

The authors and colleagues [9] developed an alternative quantitative approach to solution V&V specifically for already developed CFD codes for industrial applications (geometry and domain; models; initial, boundary and other conditions; fluid properties) and input parameters (such as iteration numbers and grid and time step sizes), which differs considerably from previous approaches. It is assumed that code verification and quality assurance issues have already been dealt with during code development. Similarly, if appropriate, it is assumed that model validation for simplified problems has also already been dealt with during model development. The philosophy is strongly influenced by EFD uncertainty analysis [10], including use of EFD definitions for errors and uncertainties. The methodology is based on concepts, definitions, and equations derived for simulation errors and uncertainties, which provide the overall mathematical framework. Verification procedures for estimating numerical errors and uncertainties include (1) the options of estimating the numerical uncertainty or the numerical error itself, which is used to obtain a corrected solution, and its uncertainty; and

(2) the concept of correction factors based on analytical benchmarks. Previously developed validation methodology and procedures for estimating modelling errors and uncertainties [11] were extended to include the option of use of corrected solutions.

The V&V approach [9] has been shown to be successful in establishing intervals of V&V for RANS simulations for ship hydrodynamics by present authors and colleagues [12–19] and international colleagues through its use at Gothenburg 2000 Workshop on CFD in Ship Hydrodynamics [20] and CFD Workshop Tokyo 2005 [21].

A shortcoming of this and other V&V approaches is that the justification for uncertainty estimates at 95% confidence level is based on reasoning similar to that used for EFD bias uncertainties at the 0-order-replication level without additional statistical 1-order-replication level precision uncertainties, which in combination provides N-order-replication level and increased confidence for EFD uncertainty analysis. Recently, present authors addressed this issue through development of a statistical approach for CFD code certification [22]. As with V&V there are many viewpoints on certification [2]. Certification is defined as a process for assessing probabilistic confidence intervals for CFD codes for specific benchmark applications and certification variables. Presumably, range of applications requires interpolation and extrapolation methods. The approach combines previous V&V approach with extensions of concept of N-version testing [23] for consideration bias uncertainties and use of reference values (experimental data and uncertainties) for estimating interval of certification. The authors have developed a similar approach for estimating intervals of certification or biases of EFD facilities or measurement systems including uncertainties [24].

The V&V approach [9] has also undergone criticism; most notably, Oberkampf [25] believes there is a conceptual flaw in the proposed approach to validation based on 3 criticisms and Roache [26] presents seven criticisms of the correction-factor verification method. We disagree with Oberkampf [25], as discussed in Reference [27]. We agree with the first criticism of Roache [26] and made appropriate revisions of the correction factor uncertainty estimates; however, we disagree with the other criticisms, as discussed in Reference [28].

The present paper summarizes and combines the V&V of CFD simulation and certification of CFD code approaches into a single presentation. Clarification is provided for some important aspects. An example of the application of quantitative certification of RANS codes is presented for ship hydrodynamics. Opportunities and challenges for achieving consensus and standard V&V and certification methodology and procedures are discussed.

SIMULATION AND CODE LEVEL ERRORS AND UNCERTAINTIES

The simulation error δ_S is defined as the difference between a simulation result S and the truth T (objective reality) and is assumed composed of additive modelling δ_{SM} and numerical δ_{SN} errors

$$\delta_S = S - T = \delta_{SM} + \delta_{SN} \tag{1}$$

Modelling errors are due to the mathematical physics problem formulation in terms of a continuous initial boundary value problem (IBVP), whereas numerical errors are due to numerical solution of the discrete IBVP. Equation (1) is considered a reasonable first approximation in consideration of the development and execution of CFD codes; however, correlations between modelling and numerical errors are also possible and should be considered in the future.

Appendix A of Reference [29] provides a derivation of (1) based on linear operator theory, which clarifies the source and role of modelling and numerical errors and provides additional justification for their being additive. The simulation uncertainty equation follows directly by considering (1) as a data reduction equation, as per EFD uncertainty analysis

$$U_S^2 = U_{SM}^2 + U_{SN}^2 \tag{2}$$

For simulations (unlike experiments except for calibrations), it is possible under certain conditions to estimate the numerical error both in sign and magnitude δ_{SN}^* such that

$$\delta_{SN} = \delta_{SN}^* + \varepsilon_{SN} \tag{3}$$

where ε_{SN} is the error in the estimate. In this case, the simulation value is corrected to provide a numerical benchmark S_C, which is defined as

$$S_C = S - \delta_{SN}^* \tag{4}$$

with corrected simulation error δ_{S_C} and uncertainty U_{S_C} equations

$$\delta_{S_C} = S_C - T = \delta_{SM} + \varepsilon_{SN} \tag{5}$$

$$U_{S_C}^2 = U_{SM}^2 + U_{S_C N}^2 \tag{6}$$

where $U_{S_C N}$ is the uncertainty estimate for ε_{SN}. The concept of deterministic error estimate for simulations seems appropriate and has been advocated by others. Equations (1)–(6) are fundamental to the present V&V of CFD simulation and certification of CFD code approaches and form the basis of the definitions, methodology, and procedures that follow.

V&V of CFD simulations are conducted at the individual user, code, model, grid-type, etc. level (simulation level) for specified applications with available benchmark EFD validation data and uncertainties. The error estimates are based on fixed values with same reasoning for 95% confidence level as used for single-realizations and 0-order-replication level bias limit estimation in EFD. This is the smallest uncertainty interval that could be achieved with the CFD code for the specified application.

Certification of CFD codes is done at the multiple codes or users, models, grid-types, etc. level (code level), which has been refereed to as N-version testing [23]. Multiple users are appropriate for code development and many user codes, whereas multiple codes is appropriate for various few user codes (often case for industrial applications). Differences between versions and implementations are due to myriad possibilities for modelling, numerical methods, and their implementation as CFD codes and simulations. As with estimating EFD precision limits using multiple realizations and 1-order-replication level testing, such estimates only include those factors turned on, which can be used to isolate differences, e.g. by using same models or grid types. In all cases, 95% confidence uncertainty estimates are based on normal distributions with similar reasoning as used for EFD precision uncertainties.

Estimating uncertainties in simulation results from CFD codes is essential when using such results in making design decisions just as is the case in using EFD results. Thus, just as EFD uncertainty analysis is required to have measurable confidence in an experiment, CFD uncertainty analysis is required to have measurable confidence in a simulation.

VERIFICATION AND VALIDATION OF CFD SIMULATIONS

Verification methodology

Verification is defined as a process for assessing simulation numerical uncertainty U_{SN} and, when conditions permit, estimating the sign and magnitude δ^*_{SN} of the simulation numerical error itself and the uncertainty in that error estimate $U_{S_C N}$.

The numerical error is decomposed into contributions from iteration number δ_I, grid size δ_G, time step δ_T, and other input parameters δ_P

$$\delta_{SN} = \delta_I + \delta_G + \delta_T + \delta_P = \delta_I + \sum_{j=1}^{J} \delta_j \tag{7}$$

which again as first approximation errors are assumed additive with simulation numerical uncertainty

$$U^2_{SN} = U^2_I + U^2_G + U^2_T + U^2_P \tag{8}$$

Similarly, for the corrected simulation

$$\delta^*_{SN} = \delta^*_I + \delta^*_G + \delta^*_T + \delta^*_P \tag{9}$$

$$U^2_{S_C N} = U^2_{I_C} + U^2_{G_C} + U^2_{T_C} + U^2_{P_C} \tag{10}$$

Substituting (9) into (4) yields

$$S_C = S - \left(\delta^*_I + \sum_{j=1}^{J} \delta^*_j \right) \tag{11}$$

or

$$S = S_C + \left(\delta^*_I + \sum_{j=1}^{J} \delta^*_j \right) \tag{12}$$

Verification procedures are based on (12).

Using (5), (11) can also be written as

$$S_C = T + \delta_{SM} + \varepsilon_{SN} \tag{13}$$

which shows that the corrected simulation result is equal to the truth plus simulation modelling error and presumable small error ε_{SN} in the estimate of the numerical error δ^*_{SN}. However, during verification the focus is on using (12) for estimating either (8) or (9) and (10), whereas as discussed next during validation focus is on estimating T and δ_{SM}.

Validation methodology and procedures

Validation is defined as a process for assessing simulation modelling uncertainty U_{SM} by using benchmark experimental data and, when conditions permit, estimating the sign and magnitude of the modelling error δ_{SM} itself.

The comparison error E is defined by the difference in the data D and simulation S values

$$E = D - S = \delta_D - (\delta_{SM} + \delta_{SN}) \tag{14}$$

where δ_D is the experimental error. Equation (14) follows directly from (1) with T estimated by $D - \delta_D$. It is assumed that D is based on an appropriate averaging of individual measurements and δ_D is estimated using standard EFD uncertainty analysis procedures as U_D. Modelling errors δ_{SM} can be decomposed into modelling assumptions and use of previous data. To determine if validation has been achieved, E is compared to the validation uncertainty U_V given by

$$U_V^2 = U_D^2 + U_{SN}^2 \tag{15}$$

If $|E| < U_V$, the combination of all the errors in D and S is smaller than U_V and validation is achieved at the U_V interval. If $U_V \ll |E|$, the sign and magnitude of $E \approx \delta_{SM}$ can be used to make modelling improvements. For the corrected simulation

$$E_C = D - S_C = \delta_D - (\delta_{SM} + \varepsilon_{SN}) \tag{16}$$

$$U_{V_C}^2 = U_{E_C}^2 - U_{SM}^2 = U_D^2 + U_{S_C N}^2 \tag{17}$$

Verification procedures

Solution verification procedures provide error and uncertainty estimates for all possible numerical error sources; however, in view of root-sum-square (RSS) representation of total uncertainty estimate only those most significant need be considered, which for many CFD solutions are iterative, grid size, and time step errors. Other error sources that may also be significant include: artificial compressibility or other similar input parameters; domain size or other similar errors; and round-off errors. For some error sources (e.g. domain size and turbulence model parameters), estimates can be based on sensitivity studies.

Procedures for estimating iterative errors are discussed in References [5, 6, 9, 30] and deserve more discussion, especially for unsteady flows.

Procedures for estimating grid size and time step errors are largely based on RE in which convergence studies are conducted with multiple, systematically refined grid sizes or time steps. RE was extended to include estimation of uncertainty using the GCI with factor of safety approach [1, 2]. GCI and other uncertainty estimates, including a new uncertainty estimate, are compared in Reference [31]. RE with concept of correction factors for estimating errors and uncertainties is used in Reference [9]. A least squares approach was proposed for estimating order of accuracy from RE when more than three grids are used and there is variability between grid studies [32]. Details of correction factor and GCI verification procedures are presented below. Wilson *et al.* [28] provides discussion and comparison of correction factor, GCI, and other approaches.

Single grid approaches to estimating numerical error are also of interest and include: (i) solution of supplemental partial differential equations for numerical error; (ii) comparing the base solution with a solution obtained with a higher-order method; and (iii) algebraic evaluations of the solution through post-processing (e.g. derivatives, conservation variables, and solution reconstruction). The first two single grid approaches provide quantitative error estimates, while the last only provides a qualitative error indicator (e.g. grid adaptation methods). Although single grid approaches have the advantage of not requiring generation of multiple grids and solutions, additional resources are required for code development, memory, and execution time associated with solution of the error equation or of the higher order solution. Studies have used the single grid approach [33, 34] and also compared single and multiple grid approaches [29, 35, 36] for simple model problems. Recent work shows promise for application of single-grid approaches to non-trivial problems [37]. In contrast to single grid approaches, the multiple grid approach can be used to establish convergence and is relatively inexpensive to implement. Other approaches such as 'manufactured solutions' (i.e. selecting an exact solution which is applied to the governing PDE to yield a source term that produces the exact solution) can be used to verify various elements of a CFD code but do not provide quantitative error estimates for a particular application.

Convergence studies. Iterative and input parameter (e.g. grid or time) convergence studies are conducted using multiple solutions (at least 3) with systematic parameter refinement by varying the kth input parameter Δx_k while holding all other parameters constant. For example, consider a time-dependent simulation with simulation numerical error given by (7) $\delta_{SN} = \delta_G + \delta_T$, assuming that iterative errors are negligible. δ_G is estimated from a grid size study using three calculations where the grid size is systematically coarsened in all coordinate directions $(\Delta x_1, \Delta x_2, \Delta x_3)$ and the time step is held constant (Δt_1). δ_T is estimated from a time step study where three calculations are performed with time step coarsening $(\Delta t_1, \Delta t_2, \Delta t_3)$ and with the grid fixed (Δx_1). The simulation with finest grid and smallest time step is common to both the grid size and time step study. Iterative errors must be accurately estimated or negligible in comparison to errors due to input parameters before accurate convergence studies can be conducted. Changes between medium–fine $\varepsilon_{k_{21}} = \hat{S}_{k_2} - \hat{S}_{k_1}$ and coarse–medium $\varepsilon_{k_{32}} = \hat{S}_{k_3} - \hat{S}_{k_2}$ solutions are used to define the convergence ratio

$$R_k = \varepsilon_{k_{21}} / \varepsilon_{k_{32}} \tag{18}$$

and to determine convergence condition where $\hat{S}_{k_1}$, $\hat{S}_{k_2}$, $\hat{S}_{k_3}$ correspond to solutions with fine, medium, and coarse input parameter, respectively, corrected for iterative errors. Four apparent convergence conditions are possible¶

$$\begin{aligned} &\text{(i) Monotonic convergence}: 0 < R_k < 1 \\ &\text{(ii) Oscillatory convergence}: R_k < 0; |R_k| < 1 \\ &\text{(iii) Monotonic divergence}: R_k > 1 \\ &\text{(iv) Oscillatory divergence}: R_k < 0; |R_k| > 1 \end{aligned} \tag{19}$$

For condition (i), generalized RE is used to estimate U_k or δ_k^* and U_{k_C}. For condition (ii), uncertainties are estimated simply by attempting to band the error based on oscillation maximums S_U and minimums S_L, i.e. $U_k = \frac{1}{2}(S_U - S_L)$, provided that the mean value is not drifting. This estimate may not be accurate if the observed values do not occur near the

¶Authors are indebted to Roache [38] for pointing out the possibility of condition (iv), which was not included in Reference [9].

minimums/maximums of the oscillations (i.e. you may need more than three samples to define the cycle and accurately estimate its uncertainty). For diverging conditions (iii) and (iv), errors and uncertainties cannot be estimated. For oscillatory convergence (ii) or divergence (iv), solutions exhibit oscillations, which may be erroneously identified as condition (i) or (iii). Oscillatory convergence issues are also discussed in References [31, 39].

Generalized RE. For convergence condition (i), generalized RE is used to estimate the error δ_k^* and order-of-accuracy p_k due to selection of the kth input parameter. The error is expanded in a power series expansion with integer powers of Δx_k as a finite sum. The accuracy of the estimates depends on how many terms are retained in the expansion, the magnitude (importance) of the higher-order terms, and the validity of the assumptions made in RE theory.

With three solutions, only the leading term can be estimated, which provides one-term estimates for error and order of accuracy

$$\delta^*_{\mathrm{RE}_{k_1}} = \frac{\varepsilon_{k_{21}}}{r_k^{p_k} - 1} \tag{20}$$

$$p_k = \frac{\ln(\varepsilon_{k_{32}}/\varepsilon_{k_{21}})}{\ln(r_k)} \tag{21}$$

Equations (20) and (21) are for uniform parameter refinement ratio r_k. For non-uniform parameter refinement ratio, Equation (20) is used with r_k replaced by $r_{k_{21}}$ and Equation (21) is replaced by

$$p_k = \frac{\ln(\varepsilon_{k_{32}}/\varepsilon_{k_{21}})}{\ln(r_{k_{21}})} - \frac{1}{\ln(r_{k_{21}})}[\ln(r_{k_{32}}^{p_k} - 1) - \ln(r_{k_{21}}^{p_k} - 1)] \tag{22}$$

In general, $m = 2n + 1$ solutions are required to estimate the first n terms of the error expansion; however, it is difficult to calculate $n > 1$ terms for practical applications since all m solutions must be both monotonically convergent and sufficiently close to asymptotic range.

There are many difficulties in applying RE to practical applications, especially demonstrating that solutions are sufficiently close to the asymptotic range. Resources for many and very fine grid size and time step are expensive. Choice of parameter refinement ratios and separation of iterative errors and grid size or time step errors for very fine values are problematic. The nature of how solutions approach the asymptotic range is not well understood except for analytical and numerical benchmarks. Nonetheless application of RE even with only three solutions can provide reasonable error and uncertainty estimates, as shown by many examples.

Correction factor C_k and GCI both use Equations (20) and (21) or (22) for estimating numerical errors δ_{SN} (7) and uncertainties U_{SN} (8), as described next. Under certain conditions, δ^*_{SN} (9) can be used to correct the simulation value and to produce a numerical benchmark S_C (4). Strictly speaking, the corrected solution will not satisfy the same conservation properties (e.g. mass and momentum) as the uncorrected solution, although if solutions are close to the asymptotic range one would expect the correction and the lack of conservation for the benchmark solution to be small. Situations that might prevent correction of the solution include variability in the observed order of accuracy, lack of complete iterative convergence, and solutions far from the asymptotic range. Situations exist when it is useful to make use of the corrected solution, e.g. for validation of verification procedures using numerical benchmarks in which the corrected solution is assumed without numerical error and modelling errors are assumed same for all solutions [40] and for practical applications as aid in assessing modelling errors when monotonic convergence established for multiple grid triplets and/or iterative convergence and resource issues for very fine grids [12, 13].

Estimating errors and uncertainties with correction factor. Correction factors provide a quantitative metric for defining distance of solutions from the asymptotic range and approximately account for the effects of higher-order terms in making error and uncertainty estimates originally based on V&V studies for 1D wave [9] and 2D Laplace equation analytical benchmarks, which showed that the one-term RE error estimate (20) has correct form, but one-term RE order-of-accuracy estimate (21) is poor except in asymptotic range. Multiplication of (20) by correction factors provides improved error and uncertainty estimates.

The numerical error is defined as

$$\delta^*_{k_1} = C_k \delta^*_{\mathrm{RE}_{k_1}} = C_k \left(\frac{\varepsilon_{k_{21}}}{r_k^{p_k} - 1} \right) \tag{23}$$

where an expression for the correction factor C_k is based on solution of (23) for C_k with $\delta^*_{\mathrm{RE}_{k_1}}$ based on (20) but replacing observed p_k with the improved estimate $p_{k_{\mathrm{est}}}$

$$C_k = \frac{r_k^{p_k} - 1}{r_k^{p_{k_{\mathrm{est}}}} - 1} \tag{24}$$

$p_{k_{\mathrm{est}}}$ is an estimate for the limiting order of accuracy of the first term as spacing size goes to zero and the asymptotic range is reached so that $C_k \to 1$. Equation (24) roughly accounts for the effects of higher-order terms by replacing p_k with $p_{k_{\mathrm{est}}}$ thereby improving the single-term estimate. Substitution of (24) into (23) results in an expression, which now requires only two solutions for evaluation

$$\delta^*_{k_1} = \left(\frac{\varepsilon_{k_{21}}}{r_k^{p_{k_{\mathrm{est}}}} - 1} \right) \tag{25}$$

However, information from all three grids is used to evaluate the uncertainty U_k, since C_k is incorporated in its definition.

The uncorrected U_k and corrected U_{k_c} solution uncertainty estimates given by Equations (33) and (34) in Reference [9] are deficient for correction factor $C_k \leqslant 1$ and $C_k = 1$, respectively, in only providing 50% uncertainty estimate (confidence level), as pointed out by Roache [26]. In Reference [13], Equation (33) in Reference [9] was revised for proper behaviour for $C_k < 1$: increasing factor of safety for decreasing C_k (i.e. distance from the asymptotic range) similarly as is the case for $C_k > 1$. Subsequently [28], revisions were also made for proper behaviour for $C_k = 1$ for both Equations (33) and (34) in Reference [9]: provision for 10% factor of safety in the limit $C_k = 1$ while smoothly merging with previous correction factor uncertainty estimates for $|1 - C_k| \geqslant 0.125$ (uncorrected solution) and 0.25 (corrected solution) given in References [9, 13]. Incorporating both revisions the uncorrected U_k and corrected U_{k_c} solution uncertainty estimates are given by

$$U_k = \begin{cases} [9.6(1-C_k)^2 + 1.1]|\delta^*_{\mathrm{RE}_{k_1}}| & |1-C_k| < 0.125 \\ [2|1-C_k| + 1]|\delta^*_{\mathrm{RE}_{k_1}}| & |1-C_k| \geqslant 0.125 \end{cases} \tag{26}$$

$$U_{k_C} = \begin{cases} [2.4(1-C_k)^2 + 0.1]|\delta^*_{\mathrm{RE}_{k_1}}| & |1-C_k| < 0.25 \\ [|1-C_k|]|\delta^*_{\mathrm{RE}_{k_1}}| & |1-C_k| \geqslant 0.25 \end{cases} \tag{27}$$

For uncorrected solutions, uncertainty estimate U_k (26) is based on the absolute value of the corrected error estimate plus the amount of the correction. For corrected solutions (i.e. corrected error estimate is used both in sign and magnitude to define numerical benchmark $S_c = S - C_k\delta^*_{\mathrm{RE}_{k_1}}$), U_{k_c} (27) is based on the absolute value of the amount of the correction.

Estimating uncertainties with GCI. In the GCI approach [1, 2] the uncertainty U_k is defined using the error estimate from RE multiplied by a factor of safety F_S to band simulation error

$$U_k = F_S|\delta^*_{\mathrm{RE}_{k_1}}| \tag{28}$$

where $\delta^*_{\mathrm{RE}_{k_1}}$ is based on a single-term estimate as given by (20) with either assumed or estimated (observed) order of accuracy. If the order of accuracy is assumed (e.g. based on theoretical values), only two solutions are required for evaluation of (20). In Reference [32], the GCI approach was extended for situations where the solution is corrected with an error estimate from RE as

$$U_{k_C} = (F_S - 1)|\delta^*_{\mathrm{RE}_{k_1}}| \tag{29}$$

An alternative expression for the uncertainty in the corrected solution has also been proposed by Equation (5.6.4) in Reference [2], which is based on a standard three-grid error estimate, i.e. (20), but with the change between fine and medium solutions $\varepsilon_{k_{21}} = S_2 - S_1$ replaced by the change between the numerical benchmark and fine solution $\varepsilon_{k_{1C}} = S_1 - S_C$ which can be written in the following alternative form:

$$U_{k_C} = \left(\frac{1}{r_k^{p_k} - 1}\right)|\delta^*_{\mathrm{RE}_{k_1}}| \tag{30}$$

This form shows that the amount of conservatism is a function of the refinement ratio r_k and order p_k, e.g. for a second-order method with $p_k = 2$, the multiplication factor applied to $|\delta^*_{\mathrm{RE}_{k_1}}|$ is 2.4, 1, $\frac{1}{3}$ for $r_k = \sqrt[4]{2}, \sqrt{2}, 2$, respectively. Thus for $r_k < \sqrt{2}$, the multiplication factor is greater than 1.0 giving the unacceptable result that the uncertainty in the corrected solution is larger than the error in the uncorrected solution. The term 'factor of safety' is used in analogy to its use in design problems. In a design environment the particular value for the factor of safety is influenced by many factors including intended purpose, conditions, and level of confidence in the analysis. For the GCI, the value for factor of safety is based on empirical data and $F_S = 1.25$ is recommended in Reference [2] for careful grid studies and $F_S = 3.0$ for cases in which only two grids are used and order of accuracy is assumed from the theoretical value p_{th}.

CERTIFICATION OF CFD CODES

At code level uncertainties are further decomposed into systematic (bias) and random (precision) components

$$U_S^2 = B_S^2 + P_S^2 = (B_{SM}^2 + B_{SN}^2) + P_S^2 \tag{31}$$

where bias uncertainties are estimated at the simulation (single realization) level and precision uncertainties at the code (N-version, multiple realization) level. Equation (31) can be written both for an individual code S_i and the average of N-version codes (mean code)

$$\bar{S} = \frac{1}{N}\sum_{i=1}^{N} S_i \tag{32}$$

as

$$U_{S_i}^2 = B_{S_i}^2 + P_{S_i}^2 \tag{33}$$

and

$$U_{\bar{S}}^2 = B_{\bar{S}}^2 + P_{\bar{S}}^2 \tag{34}$$

V&V studies (simulation level) provide

$$B_{S_i}^2 = B_{SM_i}^2 + B_{SN_i}^2 \tag{35}$$

where modelling uncertainties are decomposed into known (e.g. use of previous data or from component analysis) and estimated and unknown (e.g. due to modelling assumptions)

$$B_{SM_i}^2 = B_{SME_i}^2 + B_{SMA_i}^2 \tag{36}$$

and numerical uncertainties are decomposed into estimates for iterative, grid, time, and other input parameters

$$B_{SN_i}^2 = B_{SNI_i}^2 + B_{SNG_i}^2 + B_{SNT_i}^2 + B_{SNP_i}^2 \tag{37}$$

The mean code bias error is based on the average RSS for the individual codes

$$B_{\bar{S}}^2 = \frac{1}{N}\sum_{i=1}^{N} B_{S_i}^2 = B_{\overline{SME}}^2 + B_{\overline{SMA}}^2 + B_{\overline{SN}}^2 \tag{38}$$

N-version testing (code level) provides

$$P_{S_i} = 2\sigma_{S_i} \tag{39}$$

where

$$\sigma_{S_i} = \left[\frac{1}{N-1}\sum_{i=1}^{N}(S_i - \bar{S})^2\right]^{1/2} \tag{40}$$

and

$$P_{\bar{S}} = \frac{2\sigma_{S_i}}{\sqrt{N}} \tag{41}$$

For normal S_i distributions, the estimated truth $\bar{S}_{\mathrm{ET}}$ lies within the confidence intervals

$$S_i - U_{S_i} \leqslant \bar{S}_{\mathrm{ET}} \leqslant S_i + U_{S_i} \tag{42}$$

and

$$\bar{S} - U_{\bar{S}} \leqslant \bar{S}_{\mathrm{ET}} \leqslant \bar{S} + U_{\bar{S}} \tag{43}$$

Assume $N \geqslant 10$ and codes sufficiently similar in modelling and numerical methods and code development that S_i distribution is approximately normal. Such distributions are realizable both in ship hydrodynamics and aerodynamics for certain applications; however, multiple peaks and non-normal distributions are also realized and should be expected, e.g. clustering around turbulence models or grid types.

The mean comparison error is defined as the difference between the mean code and experimental values

$$E = D - \bar{S} \tag{44}$$

with uncertainty

$$U_E^2 = U_D^2 + U_{\bar{S}}^2 \tag{45}$$

where from (34) and (38)

$$U_{\bar{S}}^2 = B_{\overline{SME}}^2 + B_{\overline{SMA}}^2 + B_{\overline{SN}}^2 + P_{\bar{S}}^2 \tag{46}$$

Following the same reasoning and approach used for validation, certification uncertainty U_C is defined as

$$U_C^2 = U_E^2 - B_{\overline{SMA}}^2 = U_D^2 + B_{\overline{SME}}^2 + B_{\overline{SN}}^2 + P_{\bar{S}}^2 \tag{47}$$

and for

$$|E| \leqslant U_C \tag{48}$$

mean code is certified at interval U_C, whereas for

$$|E| > U_C \tag{49}$$

mean code is not certified due to modelling assumptions. E can be used for modelling assumptions improvements. In particular, for

$$|E| \gg U_C \tag{50}$$

$$E \approx \delta_{\overline{SMA}} \tag{51}$$

Similar analysis can be done for the individual code

$$E_i = D - S_i \tag{52}$$

$$U_{E_i}^2 = U_D^2 + U_{S_i}^2 \tag{53}$$

$$U_{S_i}^2 = B_{SME_i}^2 + B_{SMA_i}^2 + B_{SN_i}^2 + P_{S_i}^2 \tag{54}$$

$$U_{C_i}^2 = U_{E_i}^2 - U_{SMA_i}^2 = U_D^2 + B_{SME_i}^2 + B_{SN_i}^2 + P_{S_i}^2 \tag{55}$$

and for

$$|E_i| \leqslant U_{C_i} \tag{56}$$

individual code is certified at interval U_{C_i}, whereas for

$$|E_i| > U_{C_i} \tag{57}$$

individual code is not certified due to modelling assumptions. E_i can be used for modelling assumptions improvements. In particular, for

$$|E_i| \gg U_{C_i} \tag{58}$$

$$E_i \approx \delta_{SMA_i} \tag{59}$$

Equations (44)–(59) are similar to References [9, 11] equations for validation, except therein U_{S_i} neglects P_{S_i} such that

$$U_{C_i}^2 = U_{V_i}^2 = U_D^2 + B_{SME_i}^2 + B_{SN_i}^2 \tag{60}$$

the validation uncertainty. Certification provides additional confidence compared to validation; since, additionally based on statistics of normal distribution of N-versions. Certification uncertainty is also an improvement over simply identifying outliers [23] based on P_{S_i} alone; since additionally it includes considerations of bias uncertainties. As with EFD uncertainty analysis, maximum confidence is achieved if both bias and precision uncertainties are considered. Reference [22] provides subgroup analysis procedures for isolating and assessing differences due to use of different models and/or numerical methods, including comparisons with the method of analysis of the means [23].

EXAMPLE CERTIFICATION OF RANS CODES FOR SHIP HYDRODYNAMICS

An example is provided for certification of RANS CFD codes for ship hydrodynamics based on the Gothenburg 2000 Workshop on CFD in Ship Hydrodynamics [20]. Twenty participating research groups from 11 countries and 16 different RANS codes were used for simulations of 3 test cases representing tanker (13 submissions), container (7 submissions), and surface combatant (7 submissions) hull forms for which benchmark experimental data and uncertainties were available. The results for the tanker test case resistance (including friction and pressure components) indicated normal distributions and are used as example of certification. The tanker (KVLCC2) test case is a modern hull form with bulbous bow and stern. In this case, the free surface was not included in the simulations; since, the design Froude number (Fr) = 0.142 and wave effects small. The results for the container (Fr = 0.25) and surface combatant (Fr = 0.28) test case resistance did not indicate normal distributions due possibly to fewer submissions and complexity of free surface modelling. For KVLCC2, there were 13 submissions using 9 codes with 1 zero equation (Baldwin–Lomax), 2 one-equation (Spalart–Allmaras), 7 two-equation (4 k-w and 3 k-e), and 3 Reynolds stress turbulence models. Table I lists the organizations, codes, turbulence models, and summary of V&V and certification results for individual codes. Table II summarizes V&V and certification results for the mean code. Figure 1 shows the histogram and running record using the mean and standard deviation. Reference [22] includes results using median and median absolute deviation and subgroup analysis for 1 and 2 equation turbulence models.

Verification

The workshop recommended that the verification procedures proposed by the 22nd ITTC Resistance Committee [41] be used, which are largely based on Reference [9]. Estimates from alternative methods, including extrapolation of uncertainty estimates from benchmark applications, was also acceptable. Most groups implemented the recommended procedures, but lack of familiarity and use of coarser grids with solutions far from the asymptotic range led to difficulties. Results from grid studies showed a mixture of monotonic convergence with orders of accuracy in some cases far from expected values, oscillatory convergence, and even divergence. In spite of the difficulties, the effort was beneficial in quantitative evaluation of intervals of V&V. Quantitative estimates for iterative uncertainty U_I [B_{SNI_i} in Equation (37)] were usually not given; however, many groups mention that it is small compared to grid uncertainty, e.g. about 0.1%S (where S is solution on finest grid). For KVLCC2 the grid density varies from 0.2 to 7.9 M, but most, including those providing the most detailed verification, use about 1 M (finest grid). Five groups show convergence for the total resistance C_T with grid uncertainty U_G [B_{SNG_i} in Equation (37)] ranging from 0.1 to 4.8%$\bar{S}$ with average value $U_{G_{\text{ave}}} = 2.9\%\bar{S}$, as shown in Table I. Excluding the oscillatory convergence case, which is an order of magnitude smaller than the others, the results seem reasonable $U_{G_{\text{ave}}} = 3.6\%\bar{S}$ in comparison with values reported for the other two hull forms: $U_{G_{\text{ave}}} = 2.3$ and 5.8%$\bar{S}$ for container and surface combatant hull forms, respectively. Other submissions listed in Table I did not provide quantitative uncertainty estimates for U_G.

Validation

At the workshop, focus of validation was on prediction of mean and turbulent wake based on local flow wind tunnel measurements for a double body (Fr = 0.0) and, unfortunately, no resistance data are available. Experimental value used for example is based on ITTC model-ship correlation line and average form factor k predicted by CFD codes [$D = (1+k)C_{F0} = 4.302 \times 10^{-3}$ where $C_{F0} = 3.45 \times 10^{-3}$ and $k = 0.247$] with uncertainty $U_D = 2.2\%D$ based on best estimate for towing tank resistance test. The purpose is mainly to display proposed procedures, but values are reasonable approximations based on other hulls. The comparison error for the mean code E(44) is small $-0.1\%\bar{S}$, whereas for the individual codes E_i(52) ranges from $\pm 10\%\bar{S}$, as shown in Table I. For the codes (simulations) for which numerical uncertainties are available, the comparison error ranges from 0.6 to 5%$\bar{S}$ and validation uncertainty U_{V_i}(60) ranges from 2.2 to 5.3%$\bar{S}$ with average value 3.8%$\bar{S}$ such that only the code

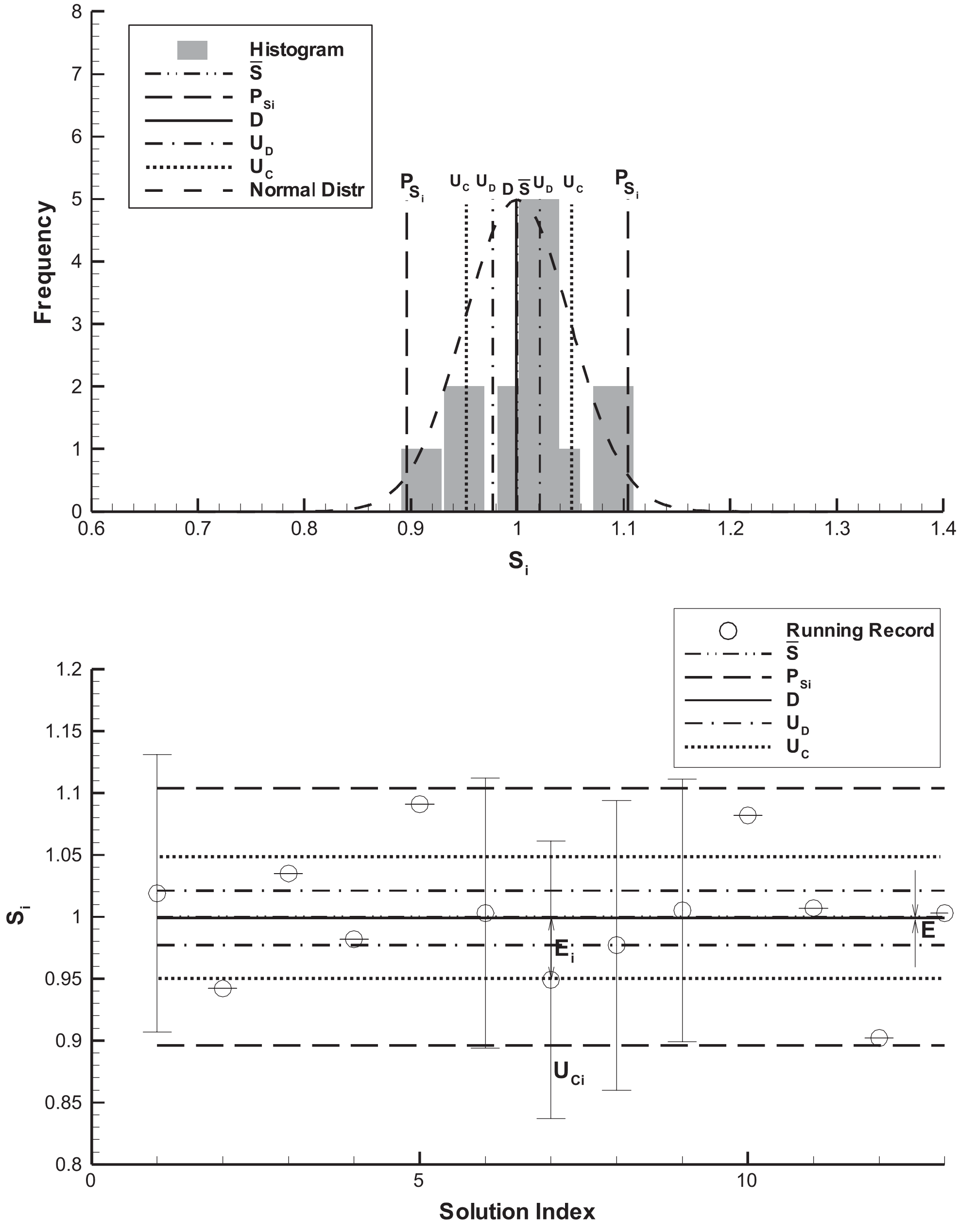

Figure 1. Histogram, normal distribution and running record for the KVLCC2 test case using mean and standard deviation.

with the largest comparison error $5\%\bar{S}$ is not validated. Other than the oscillatory convergence case, the simulation numerical uncertainty is much larger than experimental uncertainty. The mean code is validated at $3.9\%\bar{S}$ (Table II).

Table I. List of organizations, codes, turbulence models, and summary of V&V and certification results for individual codes.

								Using mean					Using median							
Index	Organization	Code	Turbulence Model	C_T (10^{-3})	C_F (10^{-3})	C_P (10^{-3})	k	S_i	D	E_i (%)	U_{c_i} (%)	P_{s_i} (%)	S_i	D	E_i (%)	U_{c_i} (%)	P_{s_i} (%)	U_V (%)	U_D (%)	B_{SN_i} (%)
1	CTU	FLUENT	k-w	4.392	3.441	0.951	0.273	1.019	0.999	−2.0	11.16	10.4	1.016	0.995	−2.1	8.44	7.4	4.05	2.2	3.4
2	FLUENT	FLUENT	RSM	4.059	3.357	0.702	0.177	0.942	0.999	5.7		10.4	0.939	0.995	5.6		7.4		2.2	
3	ECN	HOURS-easm	RSM	4.460	3.630	0.830	0.293	1.035	0.999	−3.6		10.4	1.032	0.995	−3.7		7.4		2.2	
4	ECN	HOURS-rij-w	ARS	4.230	3.300	0.930	0.226	0.982	0.999	1.7		10.4	0.978	0.995	1.7		7.4		2.2	
5	ECN	HOURS-sst	k-w	4.700	3.920	0.780	0.362	1.091	0.999	−9.2		10.4	1.087	0.995	*−9.2*		7.4		2.2	
6	MARIN-IST	PARNASSOS	k-w	4.323	3.870	0.453	0.253	1.003	0.999	−0.4	10.94	10.4	1.000	0.995	−0.5	8.15	7.4	3.41	2.2	2.6
7	SRI	NEPTUNE	SA	4.090	3.320	0.772	0.185	0.949	0.999	5.0	11.19	10.4	0.946	0.995	4.9	8.48	7.4	4.13	2.2	3.5
8	SRI	SURF	SA	4.210	3.370	0.844	0.220	0.977	0.999	2.2	11.66	10.4	0.974	0.995	2.1	9.09	7.4	5.28	2.2	4.8
9	SVA-AEA	CFX	k-w	4.329	3.397	0.932	0.255	1.005	0.999	−0.6	10.63	10.4	1.001	0.995	−0.6	7.72	7.4	2.20	2.2	0.1*
10	SOTON	CFX	k-e	4.660	3.590	1.060	0.351	1.082	0.999	−8.3		10.4	1.078	0.995	*−8.3*		7.4		2.2	
11	USDDC	UVW	BL	4.340	2.910	1.430	0.209	1.007	0.999	−0.8		10.4	1.004	0.995	−0.9		7.4		2.2	
12	KRISO	WAVIS	k-e	3.886	3.361	0.525	0.156	0.902	0.999	9.7		10.4	0.899	0.995	*9.6*		7.4		2.2	
13	MSU	UNCLE	k-e	4.320	3.560	0.760	0.252	1.003	0.999	−0.4		10.4	0.999	0.995	−0.4		7.4		2.2	
			AVE	4.308	3.464	0.844	0.247	1.000	0.999	−0.1	11.12	10.4						3.81	2.2	2.88
			STD (%)	5.2	7.5	29	25													
			Median	4.323	3.397	0.830	0.252						1.000	0.995	−0.5	8.44	7.4	4.05	2.2	3.4
			MAD (%)	4.0	4.4	19	23													

All percentages based on their mean or median values.

*Oscillatory convergence; Italic: outliers.

Table II. Summary of V&V and certification results for mean code.

	$\bar{S}$	D	E (%)	U_C (%)	U_V (%)	U_D (%)	$B_{\overline{SN}}$ (%)	$P_{\bar{S}}$ (%)
Using mean	1.000	0.999	−0.1	4.9	3.90	2.20	3.27	2.88
Using median	1.000	0.995	−0.5	4.4	3.90	2.20	3.27	2.05

All percentages based on $\bar{S}$.

Certification

Certification is done using both previous definitions of mean and standard deviation and corresponding definitions of median and median absolute deviation (MAD), which provide more robust estimates [23]. For the codes (simulations) for which numerical uncertainties are available, the certification uncertainty $U_{C_i}(55)$ ranges from 10.6 to 11.7%$\bar{S}$, which is dominated by the individual code precision uncertainty, as shown in Table I. All codes are certified at an average interval of 11.1%$\bar{S}$. Mean code, however, is certified at interval of 4.9%$\bar{S}$, which is dominated by simulation numerical bias uncertainty (Table II). Results using median and MAD are not shown but similar with reduction in values: average interval of 8.4%$\bar{S}$ and mean code 4.4%$\bar{S}$. Results are also shown graphically in Figure 1. Based on P_{S_i} alone, 0/3 outliers are identified using mean/median, respectively.

Discussion

V&V alone indicates that the various code simulations are validated at about 4%$\bar{S}$ interval. Reduction interval primarily requires reduction in simulation numerical uncertainty since it is much larger than EFD data uncertainty. Additionally, certification indicates codes certified at about 11%$\bar{S}$ interval. Reduction interval primarily requires reduction in the scatter in the results from the various codes. The overall conclusion is to reduce the intervals of verification; since, this will reduce the intervals of validation and hopefully differences between the codes and thereby also the intervals of certification. Present scatter in results (intervals P_{S_i}) is similar to intervals for aerodynamics application for wing-body configuration [23].

DISCUSSION STANDARDIZATION OF METHODOLOGY AND PROCEDURES

The solid opportunity for achieving consensus and standard V&V and certification methodology and procedures rests in the rapid increase of use of CFD for design, which absolutely requires quantitative estimation of intervals of uncertainty and levels of confidence. At the same time, challenges are seemingly insurmountable, as many viewpoints exist on nearly all aspects of CFD uncertainty analysis/accuracy estimation from basic definitions and equations (or lack of equations) to methodology to detailed procedures. Interestingly, largest differences are not with actual procedures, but with definitions and methodology, as discussed next specifically for V&V for CFD simulations, which may offer a ray of hope.

Most agree that solution verification procedures involve assessing numerical errors, especially iterative, grid size, and time step errors. Although there are some differences in detailed verification procedures for assessing iterative and grid or time convergence, e.g. use of residuals vs dependent variable time histories and GCI vs correction factors, and there is a continuing need for identification of additional error sources and improved estimation methods these hardly seem *show stoppers*. The differences arise in how one makes use of the numerical error estimates such as in estimating numerical accuracy vs numerical uncertainties and in how those estimates are used. AIAA [3, 4] estimates numerical accuracy, whereas ASME [7], Roache [2, 42], and present approach estimate numerical uncertainties. In all but present approach estimates stand alone, i.e. equations are not provided for combining such estimates for an overall estimate of numerical uncertainty/accuracy. Similarly in all but present approach estimates are end in themselves and do not play an explicit role in validation.

Similarly, most agree that validation procedures involve assessing modelling errors by comparing CFD solutions with EFD data, including consideration of EFD uncertainty. The differences are in role of numerical errors and validation quantification. AIAA [3, 4] approach quantifies validation based on percent error with caveat that numerical accuracies and experimental uncertainties should not be too large. Present approach quantifies interval of validation by comparison of percent error with RSS of all estimated numerical and experimental uncertainties. Recently, Roache [42] retracted his previous criticism [43] of [11] and is 'now in agreement with [11].'

These differences in V&V procedures arise from differences in definitions for simulation errors and uncertainties and V&V methodology, as discussed in the Introduction, which to some extent can be considered semantics, but not completely. Most important issues seem to be definitions used for errors and uncertainties and focus on these or accuracies; use of mathematical framework vs lists of activities for defining V&V; and determining level of confidence for error and uncertainty or accuracy estimates.

AIAA definitions for error and uncertainty are broad and as already mentioned based on information theory. EFD definitions used by Roache [2, 42], ASME [7], and present are not contradictory with AIAA [3, 4], but are more precise for CFD applications. AIAA [3, 4] emphasize on estimating accuracy vs errors and uncertainties is largely semantics; since, error and accuracy are inextricable and whichever used requires uncertainty estimate and confidence level for practical application.

Roache [2], AIAA [3, 4], and ASME [7] provide broad definitions, phrases and lists of activities as definitions of V&V, whereas present approach provides equations, which clearly demarcate verification as dealing with estimation of numerical uncertainty and validation with estimation of modelling uncertainty. There is no ambiguity in these definitions, whereas aforementioned phrases and activities can be ambiguous, e.g. when using phrases for verification, the authors suggest a further clarification by adding the word 'numerical analysis' (i.e. 'verification deals with mathematics of numerical analysis'); since, mathematics is also involved in modelling. Similarly, defining verification with regard to reducing errors in a computational model and validation with regard to quantification of numerical error are mixture of verification and validation. Reference [25] levelled this same criticism against present approach; however, we disagree with this reasoning since it involves substitution of D for T in Equation (13) with the claim that this implies our discussion of verification is in terms of experimental measurements. The authors do not advocate such a substitution nor focus on Equation (13) during verification, but rather on (12), which in no way involves experimental data or a mixture of verification and validation. In fact, all present V&V equations follow directly from Equation (1) and only insofar as Equation (1) obviously involves both modelling and numerical errors can one argue that we have mixed verification and validation. Reference [25] also criticized present validation approach; since, as numerical and experimental uncertainties increase, the interval of validation increases and it is easier to obtain validation (similarly as Reference [43] and rebutted by Coleman and Stern [44]). Although true, this hardly seems beneficial since a large interval of validation is clearly not useful from the design point of view just as it may be easier to conduct an experiment with large uncertainty, but usefulness of resulting data is limited. Purpose of validation is to assess interval of modelling uncertainty and thereby ascertain usefulness of modelling approach. From a design point of view it is essential to estimate intervals of V&V so as to distinguish with confidence between good and bad designs.

Uncertainty estimates require an assessment of corresponding level of confidence such as 95, 50%, etc. In EFD bias errors are estimated with 95% confidence based on assumption that the systematic error for a given error source is a fixed, single realization from a random distribution of possible errors and precision errors are estimated at 95% confidence based on statistical analysis of normal distributions procured through replication testing [10]. V&V of CFD simulations are single realization uncertainty estimates, which is basis of our interpretation of those errors as biases, whereas certification of CFD codes are multiple realization uncertainty estimates, which is basis of our interpretation of those errors as stochastic. In both cases arguments for 95% confidence are directly analogous to EFD. Roache [42] suggests an ensemble of problems in References [2, 40] provide statistical evidence for establishing 95% confidence level. We find no statistical distributions in Reference [2] or [40] while Reference [40] assumes normal distributions with the claim that twice the standard deviation leads to a 95% confidence level with no supporting evidence. Truncation errors are systematic errors with strong spatial and temporal correlations; therefore, we do not expect errors for single problems (i.e. individual user, code, model, grid-type, etc.) to display normal distributions. As with experimental uncertainty analysis, issues of replication level must be considered.

CONCLUDING REMARKS

The words of wisdom from Aristotle 'To give a satisfactory decision as to the truth it is necessary to be rather an arbitrator than a party to the dispute' and Einstein 'Truth is what stands the test of experience' seem apropos both of which place the burden on CFD users for more rigorous and widespread implementation V&V; since, currently many users ignore or only partially implement V&V. Lack of resources and standard procedures can be cited as justification, but are not really defensible if quality assurance is important.

ACKNOWLEDGEMENTS

This research was sponsored by the Office of Naval Research under Grant N00014-01-1-0073, under the administration of Dr Patrick Purtell.

REFERENCES

1. Roache P. Perspective: a method for uniform reporting of grid refinement studies. *Journal of Fluids Engineering* (ASME) 1994; **116**:405–413.
2. Roache PJ. *Verification and Validation in Computational Science and Engineering*. Hermosa Publishers: Albuquerque, NM, 1998.
3. Rahaim CP, Oberkampf WL, Cosner RR, Dominik DF. AIAA committee on standards for CFD-status and plans. *AIAA-2003-0844*, *41st Aerospace Sciences Meeting*, Reno, Nevada, 6–9 January 2003.

4. AIAA. Guide for the verification and validation of computational fluid dynamics simulations. *AIAA-G-077-1998*.
5. Roy CJ. Verification of codes and solutions in computational simulation. *Proceedings CHT-04*, Norway, 19–24 April 2004.
6. Roy CJ, Blottner FG. Assessment of one- and two-equation turbulence models for hypersonic transitional flows. *Journal of Spacecraft and Rockets* 2003; **40**(3):313–325.
7. Freitas CJ, Ghia U, Celik I, Roache P, Raad P. ASME's quest to quantify numerical uncertainty. *41st Aerospace Sciences Meeting*, Reno, Nevada, 6–9 January 2003.
8. Casey M, Wintergeste T (eds). ERCOFTAC special interest group on quality and trust in industrial CFD—best practice guidelines, Version 1, published by ERCOFTAC, 2000.
9. Stern F, Wilson RV, Coleman H, Paterson E. Comprehensive approach to verification and validation of CFD simulations—Part 1: methodology and procedures. *Journal of Fluids Engineering* (ASME) 2001; **123**:793–802.
10. Coleman HW, Steele WG. *Experimentation and Uncertainty Analysis for Engineers* (2nd edn). Wiley: New York, NY, 1999.
11. Coleman HW, Stern F. Uncertainties in CFD code validation. *Journal of Fluids Engineering* (ASME) 1997; **119**:795–803.
12. Wilson RV, Stern F, Coleman H, Paterson E. Comprehensive approach to verification and validation of CFD simulations—Part 2: application for RANS simulation of a cargo/container ship. *Journal of Fluids Engineering* (ASME) 2001; **123**(4):803–810.
13. Wilson R, Stern F. Verification and validation for RANS simulation of a naval surface combatant. Standards for CFD in the aerospace industry, *AIAA 2002-0904 Aerospace Sciences Meeting*, Reno, Nevada, 14–17 January 2002.
14. Simonsen C, Stern F. Verification and validation of RANS maneuvering simulation of Esso Osaka: effects of drift, rudder angle, and propeller on forces and moments. *Computers and Fluids* 2003; **32**:1325–1356.
15. Kim J, Paterson E, Stern F. Sub-visual and acoustic modeling for ducted marine propulsor. *8th International Conference on Numerical Ship Hydrodynamics*, Busan, Korea, 22–25 September 2003.
16. Weymouth G, Wilson R, Stern F. RANS CFD prediction of pitch and heave ship motions in head seas. *Journal of Ship Research* 2005; **49**(2):80–97.
17. Xing T, Kandasamy M, Wilson R, Stern F. DES and RANS of unsteady free-surface flows. *Division for Fluid Dynamics*. *42nd AIAA Aerospace Sciences Meeting*, Reno, Nevada, 5–8 January 2004.
18. Wilson RV, Carrica PM, Stern F. Unsteady RANS method for ship motions with application to roll for a surface combatant. *Computers and Fluids*, in press.
19. Carrica PM, Wilson RV, Stern F. Unsteady RANS simulations of the ship forward speed diffraction problem. *Computers and Fluids*, in press.
20. Larsson L, Stern F, Bertram V. Benchmarking of computational fluid dynamics for ship flows: the Gothenburg 2000 workshop. *Journal of Ship Research* 2003; **47**(1):63–81.
21. Hino T. *CFD Workshop Tokyo 2005*, National Maritime Research Institute, Tokyo, Japan, 9–11 March 2005.
22. Stern F, Wilson R, Shao J. Statistical approach to CFD code certification (invited paper). *AIAA 2003-410 Applied Aerodynamics Special Session on CFD Uncertainty*, *41st Aerospace Sciences Meeting*, Reno, Nevada, 6–9 January 2003.
23. Hemsch M. Statistical analysis of CFD solutions from the drag prediction workshop. *AIAA 2002-0842*, *40th AIAA Aerospace Sciences Meeting and Exhibit*, Reno, NV, 14–17 January 2002.
24. Stern F, Olivieri A, Shao J, Longo J, Ratcliffe T. Statistical approach for estimating intervals of certification or biases of facilities or measurement systems including uncertainties. *Journal of Fluids Engineering* (ASME) 2005; **127**:604–610.
25. Oberkampf W. Discussion. *Journal of Fluids Engineering* (ASME) 2002; **124**:809–810.
26. Roache PJ. Criticisms of the correction factor verification method. *Journal of Fluids Engineering* (ASME) 2003; **125**:732–733.
27. Stern F, Wilson R. Authors closure. *Journal of Fluids Engineering* (ASME) 2002; **124**:810–811.
28. Wilson R, Shao J, Stern F. Discussion: criticisms of the 'correction factor' verification method (Roache PJ. *Journal of Fluids Engineering* (ASME) 2003; **125**:732–733). *Journal of Fluids Engineering* (ASME) 2004; **126**:704–706.
29. Stern F, Wilson RV, Coleman H, Paterson, E. Verification and validation of CFD simulations. *IIHR Report No. 407*, Iowa Institute of Hydraulic Research, The University of Iowa, September 1999.
30. Ferziger JH, Peric M. Further discussion of numerical errors in CFD. *International Journal for Numerical Methods in Fluids* 1996; **23**:1263–1274.
31. Celik I, Karatekin O. Numerical experiments on application of Richardson extrapolation with nonuniform grids. *Journal of Fluids Engineering* (ASME) 1997; **119**:584–590.
32. Eca L, Hoekstra M. On the application of verification procedures in computational fluid dynamics. *2nd MARNET Workshop*, Barcelona, Spain, 2000.
33. Shimazaki K, Himeno Y, Baba N. An attempt at evaluating numerical errors in a 2D Navier–Stokes solver. *Symposium on Quantification of Uncertainty in Computational Fluid Dynamics*, *ASME Fluids Engineering Division*, FED, vol. 158, Summer Meeting, Washington, DC, 1993; 19–28.

34. Celik I, Hu G, Badeau A. Further refinement and benchmarking of a single-grid error estimation technique. *AIAA Paper 2003-0628*, Reno, NV, 2003.
35. Ilinca C, Zhang XD, Trepanier J-Y, Camarero R. A comparison of three error estimation techniques for finite-volume solutions of compressible flows. *Computer Methods in Applied Mechanics and Engineering* 2000; **189**(4):1247–1275.
36. Zhang XD, Pelletier D, Trepanier J-Y, Camarero R. Numerical assessment of error estimators for Euler equations. *AIAA Journal* 2001; **39**(9):1706–1715.
37. Hay A, Visonneau M. Adaptive error control strategy: application to turbulent flow. *AIAA CFD Conference*, Orlando, June 2003.
38. Roache PJ. Private communication, 2002.
39. Coleman H, Stern F, Di Mascio A, Campana E. The problem with oscillatory behavior in grid convergence studies. *Journal of Fluids Engineering* (ASME) 2001; **123**:438–439.
40. Cadafalch J, Perez-Segarra CC, Consul R, Oliva A. Verification of finite volume computations on steady state fluid flow and heat transfer. *Journal of Fluids Engineering* (ASME) 2002; **124**:11–21.
41. *International Towing Tank Conference. Resistance Committee Report. Proceedings of the 22nd ITTC*, Seoul, Korea and Shanghai, China, September 1999.
42. Roache P. Errors bars for CFD. *AIAA 2003-408*, *41st Aerospace Sciences Meeting*, Reno, Nevada, 6–9 January 2003.
43. Roache P. Discussion. *Journal of Fluids Engineering* (ASME) 1998; **120**:635.
44. Coleman HW, Stern F. Authors closure. *Journal of Fluids Engineering* (ASME) 1998; **120**:635–636.

