VERIFICATION AND VALIDATION OF CFD SIMULATIONS

by

Fred Stern, Robert V. Wilson, Hugh W. Coleman*, and Eric G. Paterson

of
Iowa Institute of Hydraulic Research
and
* Propulsion Research Center
Mechanical and Aerospace Engineering Department
University of Alabama in Huntsville
Huntsville AL 35899
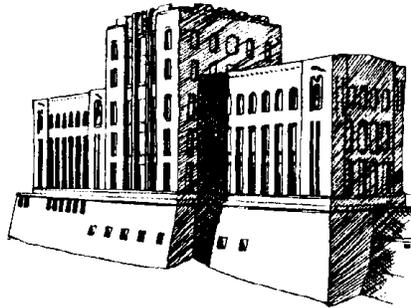
IIHR Report No. 407

Iowa Institute of Hydraulic Research
College of Engineering
The University of Iowa
Iowa City IA 52242

September 1999

Table Of Contents

Abstract

Verification and validation methodology is presented for CFD simulation results from an already developed RANS CFD code applied for specified objectives, geometry, conditions, and available benchmark information. Concepts and definitions are provided for errors and uncertainties and verification and validation. The simulation error and uncertainty equations are derived with modeling and numerical errors being additive and modeling and numerical uncertainties combining by root-sum-square. The concepts and definitions provide the mathematical framework for the verification and validation methodology.

Verification is defined as a process for assessing numerical uncertainty and, when conditions permit, estimating the sign and magnitude of the numerical error itself and the uncertainty in that error estimate. Iterative and parameter convergence studies are conducted using multiple solutions with systematic parameter refinement to estimate numerical errors and uncertainties. Three convergence conditions are possible: (i) monotonic convergence; (ii) oscillatory convergence; and (iii) divergence. For condition (i), generalized Richardson extrapolation for J input parameters and use of correction factors to account for the effects of higher-order terms and defining and estimating errors and uncertainties is used. For condition (ii), the upper and lower bounds of the solution oscillation are used to estimate uncertainties. For condition (iii), errors and uncertainties can not be estimated.

Validation is defined as a process for assessing modeling uncertainty by using benchmark experimental data and, when conditions permit, estimating the sign and magnitude of the modeling error itself. The comparison error (difference between data and simulation values) and validation uncertainty (combination of uncertainties in data and portion of simulation uncertainties that can be estimated) are used in this process.

An example is provided for a RANS CFD code and results for steady flow for a cargo/container ship.

Acknowledgements

Nomenclature

C_k	correction factor
D	benchmark data
E, E_C	comparison error, corrected
p_k	order of accuracy
R_k	parameter refinement ratio
S, S_C	simulation result, corrected
T	truth
U	uncertainty estimate
U_D	data uncertainty
U_E, U_{E_C}	comparison error uncertainty, corrected
U_I	iteration uncertainty
U_P, U_{P_C}	parameter uncertainty (e.g., grid size G and time step T), corrected
U_{reqd}	programmatic validation requirement
U_S, U_{S_C}	simulation uncertainty, corrected
U_{SM}	simulation modeling uncertainty
U_{SMA}	simulation modeling assumption uncertainty
U_{SPD}	simulation uncertainty due to use of previous data
$U_{SN}, U_{S_C N}$	simulation numerical uncertainty, corrected
U_V, U_{V_C}	validation uncertainty, corrected
Δx_k	increment in *kth* input parameter (e.g., grid size G and time step T)
d	error
d^*	error estimate with sign and magnitude
$\mathrm{d}_I, \mathrm{d}_I^*$	iteration error, estimate
$\mathrm{d}_P, \mathrm{d}_P^*$	parameter error, estimate
$\mathrm{d}_S, \mathrm{d}_{S_C}$	simulation error, corrected
d_{SN}	simulation numerical error
d_{SMA}	simulation modeling assumption error
e	solution change
e_{SN}	error in d^*

1. Introduction

Discussion and methodology for estimating errors and uncertainties in computational fluid dynamics (CFD) simulations has reached a certain level of maturity with increased attention and recent progress on common concepts and terminology (AIAA, 1998), advocacy and detailed methodology (Roache, 1998), and numerous case studies (e.g., Mehta, 1998). Progress has been accelerated in response to the urgent need for achieving consensus on concepts and terminology and useful methodology, as CFD is applied to increasingly complex geometry and physics and integrated into the engineering design process. Such consensus is required to realize the goals of simulation-based design and other uses of CFD such as simulating flows for which experiments are difficult (e.g., full-scale Reynolds numbers, hypersonic flows, off-design conditions). In spite of the progress and urgency, the various viewpoints have not converged and current methodology falls short of providing practical procedures and methodology for estimating errors and uncertainties in CFD simulations.

The present work provides a pragmatic approach for estimating errors and uncertainties in CFD simulations. Previous work on verification (Stern et al., 1996) is extended and put on a more rigorous foundation and combined with subsequent work on validation (Coleman and Stern, 1997) [hereafter referred to as C&S] thereby providing the framework for overall procedures and methodology. The philosophy is strongly influenced by experimental fluid dynamics (EFD) uncertainty analysis (Coleman and Steele, 1999), which has been standardized. Hopefully, CFD verification and validation procedures and methodology can reach a similar level of maturity and user variability can reach similar low levels, as for EFD.

The work is part of a larger program (Rood, 1996) for developing and implementing a strategy for verification and validation of Reynolds-averaged Navier-Stokes (RANS) ship hydrodynamics CFD codes. The program includes complementary CFD and EFD towing-tank investigations and considers errors and uncertainties in both the simulations and the data in assessing the success of the verification and validation efforts. The work also benefited from collaboration with the 21st and 22nd International Towing Tank Resistance Committees (ITTC, 1996 and 1999).

The focus is on verification and validation procedures and methodology for CFD simulation results from an already developed CFD code applied for specified objectives, geometry, conditions, and available benchmark information. The procedures and methodology were developed considering RANS CFD codes, but should be applicable to a fairly broad range of codes such as boundary-element methods and certain aspects of large-eddy and direct numerical simulations.

The present work differs in many respects from recent literature. The presentation is relatively succinct with intention for use for practical applications (i.e., industrial CFD) for which numerical errors and uncertainties can not be considered negligible or overlooked. The definitions of errors and uncertainties and verification and validation that are used in any approach need to be clearly stated. Table 1 summarizes the present definitions along with those given by the AIAA (1998) and Roache (1998) for comparison. The present and Roache (1998) definitions for errors and uncertainties are consistent with those used for EFD. The AIAA (1998) definitions are from an information

theory perspective and differ from those used in EFD, but are not contradictory to the present definitions. The present definitions for verification and validation are closely tied to the present definitions of errors and uncertainties and equations derived for simulation errors and uncertainties. The Roache (1998) and AIAA (1998) definitions are broader, but not contradictory to the present definitions. The present approach includes both the situations (1) of estimating errors and the uncertainty of those estimates and (2) of estimating uncertainties only. Richardson extrapolation (RE) is used for verification, which is not new; however, the present generalizations for J input parameters and use of correction factors to account for the effects of higher-order terms and in defining and estimating errors and uncertainties constitute a new approach. The use of quantitative estimates for errors and the use of uncertainties for those estimates also constitute a new approach in verification and validation.

2. Verification and Validation Procedures

The overall CFD verification and validation procedures can be conveniently grouped in four consecutive steps: (1) preparation; (2) verification; (3) validation; and (4) documentation.

Preparation. The 1st step is preparation, which involves selection of the CFD code and specification of objectives, geometry, conditions, and available benchmark information. The objectives might be prediction of certain variables at certain levels of validation (e.g., programmatic validation requirements U_{reqd}). The variables can either be integral (e.g., resistance) or point (e.g., mean velocities and turbulent Reynolds stresses) values and the programmatic validation requirements may be different for each variable.

Verification. The 2nd step is verification, which is defined as a process for assessing simulation numerical uncertainty U_{SN} and, when conditions permit, estimating the sign and magnitude d_{SN}^{*} of the simulation numerical error itself and the uncertainty in that error estimate (referred to as the corrected simulation numerical uncertainty $U_{S_C N}$). Iterative and input parameter convergence studies are conducted using multiple solutions with systematic parameter, as described in Section 3.2.

Validation. The 3rd step is validation, which is defined as a process for assessing simulation modeling uncertainty U_{SM} by using benchmark experimental data and, when conditions permit, estimating the sign and magnitude of the simulation modeling error δ_{SM} itself. The comparison error E (difference between data D and simulation S values) and validation uncertainty U_V (combination of uncertainties in data and portion of simulation uncertainties that can be estimated) are used, as described in Section 3.3.

Documentation. The 4th step is documentation, which is detailed presentation of the CFD code (equations, initial and boundary conditions, modeling, and numerical methods), objectives, geometry, conditions, verification, validation, and analysis.

3. Verification and Validation Methodology

Verification (Section 3.2) and validation (Section 3.3) methodology is presented for CFD simulation results from an already developed CFD code applied for specified objectives, geometry, conditions, and available benchmark information. Section 3.1 discusses concepts and definitions for errors and uncertainties and verification and validation, which provide the mathematical framework for the verification and validation methodology. Analytical benchmarks can be defined as the truth and are useful in development and confirmation of verification procedures and methodology and in code development, but can not be used for validation and are restricted to simple equations. Results from the use of analytical benchmarks are provided in Appendix C.

3.1 Concepts and Definitions

Accuracy indicates the closeness of agreement between a simulation/experimental value of a quantity and its true value. Error δ is the difference between a simulation value or an experimental value and the truth. Accuracy increases as error approaches zero. The true values of simulation/experimental quantities are rarely known. Thus, errors must be estimated. An uncertainty U is an estimate of an error such that the interval $\pm U$ contains the true value of δ 95 times out of 100. An uncertainty interval thus indicates the range of likely magnitudes of δ but no information about its sign.

For simulations, under certain conditions, errors can be estimated including both sign and magnitude (referred to as an error estimate δ*). Then, the uncertainty considered is that corresponding to the error in δ*. When δ^* is estimated, it can be used to obtain a corrected value of the variable of interest.

Sources of errors and uncertainties in results from simulations can be divided into two distinct sources: modeling and numerical. Modeling errors and uncertainties are due to assumptions and approximations in the mathematical representation of the physical problem (such as geometry, mathematical equation, coordinate transformation, boundary conditions, turbulence models, etc.) and incorporation of previous data (such as fluid properties) into the model. Numerical errors and uncertainties are due to numerical solution of the mathematical equations (such as discretization, artificial dissipation, incomplete iterative and grid convergence, lack of conservation of mass, momentum, and energy, internal and external boundary non-continuity, computer round-off, etc.). The present work assumes that all correlations among errors are zero, which is doubtless not true in all cases, but the effects are assumed negligible for the present analyses.

The simulation error δ_S is defined as the difference between a simulation result S and the truth T. In considering the development and execution of a CFD code, it can be postulated that δ_S is comprised of the addition of modeling and numerical errors

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

A derivation of the simulation error equation (1) is provided in Appendix A. The uncertainty equation corresponding to error equation (1) is

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

3

where U_S is the uncertainty in the simulation and U_{SM} and U_{SN} are the simulation modeling and numerical uncertainties.

For certain conditions, the numerical error d_{SN} can be considered as

$$\mathsf{d}_{SN} = \mathsf{d}^*_{SN} + \mathsf{e}_{SN} \tag{3}$$

where d^*_{SN} is an estimate of the sign and magnitude of d_{SN} and e_{SN} is the error in that estimate (and is estimated as an uncertainty since only a range bounding its magnitude and not its sign can be estimated). The corrected simulation value S_C is defined by

$$S_C = S - \mathsf{d}^*_{SN} \tag{4}$$

with error equation

$$\mathsf{d}_{S_C} = S_C - T = \mathsf{d}_{SM} + \mathsf{e}_{SN} \tag{5}$$

The uncertainty equation corresponding to error equation (5) is

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

where U_{S_C} is the uncertainty in the corrected simulation and $U_{S_C N}$ is the uncertainty estimate for e_{SN}.

Debate on verification and validation has included discussion on whether errors such as d_{SN} are deterministic vs. stochastic and thus how they should be treated in uncertainty analysis was unclear. In the approach given by equations. (3)-(6), a deterministic estimate d^*_{SN} of d_{SN} and consideration of the error e_{SN} in that estimate are used. The approach is analogous to that in EFD when an asymmetric systematic uncertainty is "zero-centered" by inclusion of a model for the systematic error in the data reduction equation and then the uncertainty considered is that associated with the model (Coleman and Steele, 1999). In the "uncorrected" approach given by equations (1)-(2), any particular d_{SN} is considered as a single realization from some parent population of d_{SN}'s and the uncertainty U_{SN} is interpreted accordingly in analogy to the estimation of uncertainties in EFD (with a similar argument for e_{SN} and $U_{S_C N}$).

3.2 Verification

For many CFD codes, the most important numerical errors and uncertainties are due to use of iterative solution methods and specification of various input parameters such as spatial and time step sizes and other parameters (e.g., artificial dissipation). The errors and uncertainties are highly dependent on the specific application (geometry and conditions).

The errors due to specification of input parameters are decomposed into error contributions from iteration number d_I, grid size d_G, time step d_T, and other parameters

d_P, which gives the following expressions for the simulation numerical error and uncertainty

$$d_{SN} = d_I + d_G + d_T + d_P = d_I + \sum_{j=1}^{J} d_j \qquad (7)$$

$$U_{SN}^2 = U_I^2 + U_G^2 + U_T^2 + U_P^2 = U_I^2 + \sum_{j=1}^{J} U_j^2 \qquad (8)$$

Similarly, error estimates d^* can be decomposed as

$$d_{SN}^* = d_I^* + \sum_{j=1}^{J} d_j^* \qquad (9)$$

which gives the following expressions for the corrected simulation and corrected simulation numerical uncertainty

$$S_C = S - (d_I^* + \sum_{j=1}^{J} d_j^*) = T + d_{SM} + e_{SN} \qquad (10)$$

$$U_{S_C N}^2 = U_{I_C}^2 + \sum_{j=1}^{J} U_{jc}^2 \qquad (11)$$

Verification is based on equation (10), which is put in the form

$$S = S_C + (d_I^* + \sum_{j=1}^{J} d_j^*) \qquad (12)$$

Equation (12) expresses S as the corrected simulation value S_C plus numerical errors. S_C is also referred to as a numerical benchmark since it is equal, as shown by equation (10), to the truth plus simulation modeling error and presumable small error e_{SN} in the estimate of the numerical error d_{SN}^*. Power-series expansions for each input parameter and multiple solutions are used to obtain estimates for the d_j^*'s in equation (12). For this approach to be useful, d_I^* must be accurately estimated or be negligible for each solution.

3.2.1 Convergence Studies

Iterative and parameter convergence studies are conducted using multiple (m) solutions and systematic parameter refinement by varying the *kth* input parameter Δx_k while holding all other parameters constant. The present work assumes input parameters can be expressed such that the finest resolution corresponds to the limit of infinitely small parameter values. Many common input parameters are of this form, e.g., grid spacing, time step, and artificial dissipation. Additionally, a uniform parameter refinement ratio $r_k = \Delta x_{k_2}/\Delta x_{k_1} = \Delta x_{k_3}/\Delta x_{k_2} = \Delta x_{k_m}/\Delta x_{k_{m-1}}$ between solutions is assumed. The use of uniform parameter refinement ratio is not required (Roache, 1998); however, it

simplifies the analysis and in the authors experience use of non-uniform parameter refinement ratio is not needed.

Careful consideration should be given to selection of uniform parameter refinement ratio. The most appropriate values for industrial CFD are not yet fully established. Small values (i.e., very close to one) are undesirable since solution changes will be small and sensitivity to input parameter may be difficult to identify compared to iterative errors. Large values alleviate this problem; however, they also may be undesirable since the finest step size may be prohibitively large if the coarsest step size is designed for sufficient resolution such that similar physics are resolved for all m solutions. Also, similarly as for small values, solution changes for the finest step size may be difficult to identify compared to iterative errors since iterative convergence is more difficult for small step size. Another issue is that for parameter refinement ratio other than $r_k = 2$, interpolation to a common location is required to compute solution changes, which introduces interpolation errors. Roache (1998) discusses methods for evaluating interpolation errors. However, for industrial CFD, $r_k = 2$ may often be too large. A good alternative may be $r_k = \sqrt{2}$, as it provides fairly large parameter refinement ratio and at least enables prolongation of the coarse-parameter solution as an initial guess for the fine-parameter solution.

Equation (12) is written for the *kth* parameter and *mth* solution as

$$S_{k_m} = S_C + d^*_{I_{k_m}} + d^*_{k_m} + \sum_{j=1, j \neq k}^{J} d^*_{j_m} \tag{13}$$

Iterative convergence must be assessed and S_{k_m} corrected for iterative errors prior to evaluation of parameter convergence since the level of iterative convergence may not be the same for all *m* solutions used in the parameter convergence studies. Methods for estimating U_I or d^*_I and U_{I_C} are described in Section 3.2.2. With $d^*_{I_{k_m}}$ evaluated, S_{k_m} is corrected for iterative errors as

$$\hat{S}_{k_m} = S_{k_m} - d^*_{I_{k_m}} = S_C + d^*_{k_m} + \sum_{j=1, j \neq k}^{J} d^*_{j_m} \tag{14}$$

Equation (13) shows that iterative errors $d^*_{I_{k_m}}$ must be accurately estimated or negligible in comparison to $d^*_{k_m}$ for accurate convergence studies and that they should be considered within the context of convergence studies for each input parameter.

S_{k_m} can be calculated for both integral (e.g., resistance coefficients) and point (e.g., surface pressure, wall-shear stress, and velocity) variables. S_{k_m} can be presented as an absolute quantity (i.e., non-normalized) or normalized with the solution as a percentage change; however, if the solution value is small, a more appropriate normalization may be the range of the solution.

Convergence studies require a minimum of *m=3* solutions to evaluate convergence with respect to input parameter. Note that *m=2* is inadequate, as it only indicates sensitivity and not convergence, and that *m>3* may be required. Consider the situation for

3 solutions corresponding to fine S_{k_1}, medium S_{k_2}, and coarse S_{k_3} values for the *kth* input parameter. Solution changes e for medium-fine and coarse-medium solutions and their ratio R_k are defined by

$$e_{21_k} = \hat{S}_{k_2} - \hat{S}_{k_1}$$

$$e_{32_k} = \hat{S}_{k_3} - \hat{S}_{k_2} \qquad (15)$$

$$R_k = e_{21_k} / e_{32_k}$$

Three convergence conditions are possible:

(i) Monotonic convergence: $0 < R_k < 1$

(ii) Oscillatory convergence: $R_k < 0^{i}$ $\qquad (16)$

(iii) Divergence: $R_k > 1$

For monotonic convergence (i), generalized RE is used to estimate U_k or d_k^* and U_{k_C}. Methods for estimating errors and uncertainties for condition (i) are described in Section 3.2.3.

For oscillatory convergence (ii), the solutions exhibit oscillations, which may be erroneously identified as condition (i) or (iii). This is apparent if one considers evaluating convergence condition from three points on a sinusoidal curve (Coleman et al., 1999). Depending on where the three points fall on the curve, the condition could be incorrectly diagnosed as either monotonic convergence or divergence. Methods discussed here for estimating uncertainties U_k for condition (ii) require more than *m*=3 solutions and are described in Section 3.2.4.

For divergence (iii), the solutions diverge and errors and uncertainties can not be estimated. Additional remarks are given in Section 3.2.5.

Determination of the convergence ratio R_k for point variables can be problematic since solution changes e_{21_k} and e_{32_k} can both go to zero (e.g., in regions where the solution contains an inflection point). In this case, the ratio becomes ill conditioned. However, the convergence ratio can be used in regions where the solution changes are both non-zero (e.g., local solution maximums or minimums). Another approach is to use a global convergence ratio R_k, which overcomes ill conditioning, based on the L2 norm of the solution changes, i.e., $\langle R_k \rangle = \|e_{21_k}\|_2 / \|e_{32_k}\|_2$. $<>$ is used to denote an averaged value and $\|e\|_2 = \left[\sum_{i=1}^{N} e_i^2 \right]^{1/2}$ denotes the L2 norm of solution change over the N points in the region of interest. Caution should be exercised when defining the convergence ratio from the ratio of the L2 norm of solution changes because the oscillatory condition ($R_k < 1$) cannot

---

[i] As discussed in the text that follows, $0 < R_k < 1$ and $R_k > 1$ may also occur for the oscillatory condition.

be diagnosed since $\langle R_k \rangle$ will always be greater than zero. Local values of R_k at solution maximums or minimums should also be examined to confirm the convergence condition based on an L2 norm definition.

3.2.2 Iterative Convergence

Iterative convergence must be assessed and simulation results S_{k_m} corrected for iterative errors prior to evaluation of parameter convergence since the level of iterative convergence may not be the same for all m solutions used in the parameter convergence studies. Methods for estimating U_I or d_I^* and U_{I_C} are described in this section. The methods are applicable to both integral and point variables. For point variables, an L2 norm over all grid points is often used as a global metric. There are many integral and point variables that can be monitored to establish iterative stopping criteria; however, present discussion is specifically within the context of evaluating U_I or d_I^* and U_{I_C} for use in the parameter convergence study for S_{k_m}. Further work is needed on assessing iterative errors and their role in parameter convergence studies and for assessing iterative errors and uncertainties for unsteady flows.

Typical CFD solution techniques for obtaining steady state solutions involve beginning with an initial guess and performing time marching or iteration until a steady state solution is achieved. For time-accurate calculations using implicit methods, convergence of the solution is required at each time step. Care must be exercised in evaluating iterative convergence based solely on solution residuals, i.e., change in solution from iteration to iteration. Small time steps and/or relaxation parameters can result in small solution residuals while iterative error can be large (Ferziger and Peric, 1997). If S_{k_m} is a primary dependent variable, an alternative approach that removes this problem is to use the residual imbalance of the discretized equations (i.e., the difference in the left- and right-hand sides) as a measure of convergence; since, the iterative error satisfies the same equation as this residual imbalance.

The number of order magnitude drop and final level of solution residual (or residual imbalance) can be used to determine stopping criteria for iterative solution techniques. Iterative convergence to machine zero is desirable, but for complex geometry and conditions it is often not possible. Three or four orders of magnitude drop in solution residual to a level of 10^{-4} is more likely for these cases. Methods for estimation of iterative errors and uncertainties can be based on graphical, as discussed below, or theoretical approaches and are dependent on the type of iterative convergence: (a) oscillatory; (b) convergent; or (c) mixed oscillatory/convergent.

For oscillatory iterative convergence (a), the deviation of the variable from its mean value provides estimates of the iterative uncertainty based on the range of the maximum S_U and minimum S_L values

$$U_I = \left| \frac{1}{2}(S_U - S_L) \right| \tag{17}$$

8

For convergent iterative convergence (b), a curve-fit of an exponential function can be used to estimate U_I or d_I^* and U_{I_C} as the difference between the value and the exponential function from a curve fit for large iteration number CF_∞

$$U_I = \left| S - CF_\infty \right|$$
$$\mathrm{d}_{I_{k_m}}^* = S - CF_\infty , U_{I_C} = 0$$

(18)

For mixed convergent/oscillatory iterative convergence (c), the amplitude of the solution envelope decreases as the iteration number increases, the solution envelope is used to define the maximum S_U and minimum S_L values in the Ith iteration, and to estimate U_I or d_I^* and U_{I_C}

$$U_I = \left| \frac{1}{2}(S_U - S_L) \right|$$
$$\mathrm{d}_{I_{k_m}}^* = S - \frac{1}{2}(S_U - S_L), U_{I_C} = 0$$

(19)

An increase in the amplitude of the solution envelope as the iteration number increases indicates that the solution is divergent.

Estimates of the iterative error based on theoretical approaches are presented in Ferziger and Peric (1997) and involve estimation of the principal eigenvalue of the iteration matrix. The approach is relatively straightforward when the eigenvalue is real and the solution is convergent. For cases in which the principal eigenvalue is complex and the solution is oscillatory or mixed, the estimation is not as straightforward and additional assumptions are required.

3.2.3 Monotonic Convergence: Generalized Richardson Extrapolation

For monotonic convergence, i.e., condition (i) in equation (16), generalized RE is used to estimate U_k or d_k^* and U_{k_C}. RE is generalized for J input parameters and use of correction factors to account for the effects of higher-order terms and defining and estimating errors and uncertainties, as summarized in the following. Appendix B provides a detailed description.

Generalized RE begins with equation (14). The error terms on the right-hand-side of equation (14) are of known form (i.e., power series expansion in Δx_k) based on analysis of the modified (A.6) and numerical error (A.9) equations, as shown in Appendix A equation (A.12), which is written below as a finite sum (i.e., error estimate) and for the *kth* parameter and *mth* solution

$$\mathrm{d}_{k_m}^* = \sum_{i=1}^{n} (\Delta x_{k_m})^{p_k^{(i)}} g_k^{(i)}$$

(20)

n = number of terms retained in the power series, powers $p_k^{(i)}$ correspond to order of accuracy (for the *ith* term), and $g_k^{(i)}$ are referred to as "grid" functions which are a

9

function of various orders and combinations of derivatives of S with respect to x_k. Substituting equation (20) into equation (14) results in

$$\hat{S}_{k_m} = S_C + \sum_{i=1}^{n}(\Delta x_{k_m})^{p_k^{(i)}} g_k^{(i)} + \sum_{j=1, j\neq k}^{J} d_{j_m}^* \tag{21}$$

Subtraction of multiple solutions where input parameter Δx_k is uniformly refined eliminates the $d_{j_m}^*$ terms in equation (21) since $d_{j_m}^*$ is independent of Δx_k and provides equations for S_C, $p_k^{(i)}$, and $g_k^{(i)}$. This assumes $p_k^{(i)}$ and $g_k^{(i)}$ are also independent of Δx_k. Since each term (i) contains 2 unknowns, $m=2n+1$ solutions are required to estimate the numerical benchmark S_C and the first n terms in the expansion in equation (21) (i.e., for $n=1$, $m=3$ and for $n=2$, $m=5$, etc). The accuracy of the estimates depends on how many terms are retained in equation (20), the magnitude (importance) of the higher-order terms, and the validity of the assumption that $p_k^{(i)}$ and $g_k^{(i)}$ are independent of Δx_k. For sufficiently small Δx_k, the solutions are in the asymptotic range such that higher–order terms are negligible and the assumption that $p_k^{(i)}$ and $g_k^{(i)}$ are independent of Δx_k is valid. However, achieving the asymptotic range for practical geometry and conditions is usually not possible and $m>3$ is undesirable from a resources point of view; therefore, methods are needed to account for effects of higher-order terms for practical application of RE. Additionally, methods may be needed to account for possible dependence of $p_k^{(i)}$ and $g_k^{(i)}$ on Δx_k, although not addressed herein. Usually d_k^* is estimated for the finest value of the input parameter, i.e., $d_k^* = d_{k_1}^*$ corresponding to the finest solution S_{k_1}.

For $m=3$, only the leading-order term can be evaluated. Equations are obtained for $d_{k_1}^*$ and order-of-accuracy p_k

$$d_{k_1}^* = d_{RE_{k_1}}^* = \frac{e_{21_k}}{r_k^{p_k} - 1} \tag{22}$$

$$p_k = \frac{ln(\varepsilon_{32_k}/\varepsilon_{21_k})}{ln(r_k)} \tag{23}$$

Appendix B includes results for $m=5$.

Appendix C provides verification for two analytical benchmarks (one-dimensional wave and two-dimensional Laplace equations). Multiple solutions were used to evaluate the RE error estimates, including the effects of higher-order terms. Solving for the first-order term is relatively easy since evaluation of equations (22) and (23) only requires that the $m=3$ solutions are monotonically convergent, even if the solutions are far from the asymptotic range and equations (22) and (23) are inaccurate. Solving for the higher-order terms (i.e., second-order term) is more difficult since evaluation of the $m=5$ solutions for S_C, $p_k^{(i=1,2)}$, and $g_k^{(i=1,2)}$ additionally requires that the solutions are relatively close to the asymptotic range, i.e., within about 6% of the theoretical order of accuracy based on the modified equation $p_{k_{th}}$ and $q_{k_{th}}$.

10

The solutions show that equation (22) has the correct form, but the order of accuracy is poorly estimated by equation (23) except in the asymptotic range. Therefore, one approach is to correct equation (22) by a multiplication correction factor to account for the effects of higher-order terms. Two correction factors were investigated

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

(24a)

$$C_k = \frac{(e_{23_k}/e_{12_k} - r_k^{q_{k\,est}})(r_k^{p_k} - 1)}{\left(r_k^{p_{k\,est}} - r_k^{q_{k\,est}}\right)\!\left(r_k^{p_{k\,est}} - 1\right)} + \frac{(e_{23_k}/e_{12_k} - r_k^{p_{k\,est}})(r_k^{p_k} - 1)}{(r_k^{p_{k\,est}} - r_k^{q_{k\,est}})(r_k^{q_{k\,est}} - 1)}$$

(24b)

$p_{k_{est}}$ and $q_{k_{est}}$ are estimates for the 1st and 2nd term order of accuracy $p_k^{(1)}$ and $p_k^{(2)}$. The estimated values can be based either on $p_{k_{th}}$ and $q_{k_{th}}$ or solutions for simplified geometry and conditions. In either case, preferably including the effects of grid stretching. Equation (24a) roughly accounts for the effects of higher-order terms by replacing p_k with $p_{k_{est}}$ thereby providing an improved single-term estimate. Equation (24b) more rigorously accounts for higher-order terms since it is derived from the two-term estimate with 1st and 2nd term order of accuracy $p_k^{(1)}$ and $p_k^{(2)}$ replaced by $p_{k_{est}}$ and $q_{k_{est}}$. Equation (24b) simplifies to equation (24a) in the limit of the asymptotic range. Both correction factors only require solutions for three parameter values. $C_k < 1$ or $C_k > 1$ indicates that the leading-order term over predicts (higher-order terms net negative) or under predicts (higher-order terms net positive) the error, respectively. C_k given by equation (24) is fairly universal in that it only implicitly depends on geometry and conditions. However, C_k is based on results from only two linear analytical benchmarks and additional benchmarks (especially non-linear) are needed to confirm the universality of equation (24) or to provide alternative forms.

Combining equation (22) and (24) provides an estimate for $d_{k_1}^*$ accounting for the effects of higher-order terms

$$d_{k_1}^* = C_k d_{RE_{k_1}}^* = C_k \left(\frac{e_{21_k}}{r_k^{p_k} - 1} \right)$$

(25)

The estimate includes both sign and magnitude. Equation (25) is used to estimate U_k or d_k^* and U_{k_C} depending on how close the solutions are to the asymptotic range (i.e., how close C_k is to 1) and one's confidence in equation (25). There are many reasons for lack of confidence, especially for complex three-dimensional flows. Point variables invariably are not uniformly convergent, which is particularly evident near inflection points and zero crossings.

Equations (24) and (25) need further testing both for additional analytical benchmarks (as already mentioned) and practical applications. Also alternative strategies for including effects of higher-order terms may be just as viable. Note that equation (25) differs

11

significantly from the GCI proposed by Roache (1998). Herein $C_k = C_k(e, r_k, p_k, p_{k\,est}, q_{k\,est})$, whereas in the GCI, C_k is a constant referred to as a factor of safety F_S which equals 1.25 for careful grid studies and 3 for cases for which only two grids are used.

For C_k sufficiently less than or greater than 1 and lacking confidence, U_k is estimated, but not d_k^* and U_{k_C}. Based on the analytical benchmark studies (Appendix C), it appears that equation (25) can be used to estimate the uncertainty by bounding the error by the sum of the absolute value of the corrected estimate from RE and the absolute value of the amount of the correction

$$U_k = \left| C_k d_{RE_{k_1}}^* \right| + \left| (1 - C_k) d_{RE_{k_1}}^* \right| \tag{26}$$

For C_k sufficiently close to 1 and having confidence, d_k^* and U_{k_C} are estimated. Equation (25) is used to estimate the error d_k^*, which can then also be used in the calculation of S_C [in equation (10)]. The uncertainty in the error estimate is based on the amount of the correction

$$U_{k_C} = \left| (1 - C_k) d_{RE_{k_1}}^* \right| \tag{27}$$

Note that in the limit of the asymptotic range, $C_k = 1$, $d_k^* = d_{k_1}^* = d_{RE_{k_1}}^*$, and $U_{k_C} = 0$.

3.2.4 Oscillatory Convergence

For oscillatory convergence, i.e., condition (ii) in equation (16), uncertainties can be estimated, but not the signs and magnitudes of the errors. Uncertainties are estimated based on determination of the upper (S_U) and lower (S_L) bounds of solution oscillation, which requires more than $m=3$ solutions. The estimate of uncertainty is based on half the solution range

$$U_k = \frac{1}{2}(S_U - S_L) \tag{28}$$

3.2.5 Divergence

For divergence, i.e., condition (iii) in equation (16), errors and or uncertainties can not be estimated. The preparation and verification steps must be reconsidered. Improvements in iterative convergence, parameter specification (e.g., grid quality), and/or CFD code may be required to achieve converging or oscillatory conditions.

3.3 Validation

Validation is defined as a process for assessing modeling uncertainty U_{SM} by using benchmark experimental data and, when conditions permit, estimating the sign and

12

magnitude of the modeling error d_{SM} itself. Thus, the errors and uncertainties in the experimental data must be considered in addition to the numerical errors and uncertainties discussed in Section 3.2. Approaches to estimating experimental uncertainties are presented and discussed by Coleman and Steele (1999).

The validation methodology of Coleman and Stern (1997) which properly takes into account the uncertainties in both the simulation and the experimental data is described in this section. The methodology is also demonstrated using an estimated numerical error and corrected simulation and validation uncertainty values.

3.3.1 Methodology

The validation comparison for a simulated and measured result r that is a function of the variable X is shown in figure 1. The experimentally determined r-value of the (X_i, r_i) **d**ata point is D and, as before, the simulated r-value is S. Recall from equation (1) that the simulation error d_S is the difference between S and the truth T. Similarly, the error d_D in the data is the difference between D and the truth T, so setting the simulation and experimental truths equal results in

$$D - d_D = S - d_S \tag{29}$$

The comparison error E is defined as the difference of D and S

$$E = D - S = d_D - d_S = d_D - (d_{SMA} + d_{SPD} + d_{SN}) \tag{30}$$

with d_{SM} decomposed into the sum of d_{SPD}, error from the use of <u>p</u>revious <u>d</u>ata such as fluid properties, and d_{SMA}, error from <u>m</u>odeling <u>a</u>ssumptions. Thus E is the resultant of <u>all</u> the errors associated both with the experimental data and with the simulation. For the approach in which no estimate d_{SN}^* of the sign and magnitude of d_{SN} is made, all of these errors are estimated with uncertainties. (As will be shown, during the validation process an estimate of the sign and magnitude of d_{SMA} can be made under certain conditions.)

If X_i, r_i, and S share no common error sources, then the uncertainty U_E in the comparison error can be expressed as

$$U_E^2 = \left(\frac{\P E}{\P D}\right)^2 U_D^2 + \left(\frac{\P E}{\P S}\right)^2 U_S^2 = U_D^2 + U_S^2 \tag{31}$$

or

$$U_E^2 = U_D^2 + U_{SMA}^2 + U_{SPD}^2 + U_{SN}^2 \tag{32}$$

where subscripts are used in the same manner as for the d's .

Ideally, we would like to postulate that if the absolute value of E is less than its uncertainty U_E, then validation is achieved (i.e., E is "zero" considering the resolution imposed by the "noise level" U_E). In reality, the authors know of no approach that gives

an estimate of U_{SMA}, so U_E cannot be estimated. That leaves a more stringent validation test as the practical alternative. If the validation uncertainty U_V is defined as the combination of all uncertainties that we know how to estimate (i.e., all but U_{SMA}), then

$$U_V^2 = U_E^2 - U_{SMA}^2 = U_D^2 + U_{SPD}^2 + U_{SN}^2 \tag{33}$$

If $|E|$ is less than the validation uncertainty U_V, the combination of all the errors in D and S is smaller than the estimated validation uncertainty and validation has been achieved at the U_V level. U_V is the key metric in the validation process. U_V is the validation "noise level" imposed by the uncertainties inherent in the data, the numerical solution, and the previous experimental data used in the simulation model. It can be argued that one cannot discriminate once $|E|$ is less than this; that is, as long as $|E|$ is less than this, one cannot evaluate the effectiveness of proposed model "improvements."

If the corrected approach of equations (3)- (6) is used, then the equations equivalent to equations (30) and (33) are

$$E_C = D - S_C = \mathsf{d}_D - (\mathsf{d}_{SMA} + \mathsf{d}_{SPD} + \mathsf{e}_{SN}) \tag{34}$$

for the corrected comparison error and

$$U_{V_C}^2 = U_{E_C}^2 - U_{SMA}^2 = U_D^2 + U_{SPD}^2 + U_{S_C N}^2 \tag{35}$$

for the corrected validation uncertainty. Note that S_C and E_C can be either larger or smaller than their counterparts S and E, but U_{E_C} and U_{V_C} should be smaller than U_E and U_V, respectively, since $U_{S_C N}$ should be smaller than U_{SN}.

For the data point (X_i, r_i), U_D should include both the experimental uncertainty in r_i and the additional uncertainties in r_i arising from experimental uncertainties in the measurements of the n independent variables $(X_j)_i$ in X_i. The expression for U_D that should be used in the U_V (U_{V_C}) calculation is then

$$U_D^2 = U_{r_i}^2 + \sum_{j=I}^{n} \left(\frac{\partial r}{\partial X_j} \right)_i^2 \left(U_{X_j} \right)_i^2 \tag{36}$$

In some cases, the terms in the summation in equation (36) may be shown to be very small, using an order-of-magnitude analysis, and then neglected. This would occur in situations in which the U_{X_j} values are of "reasonable" magnitude and gradients in r are small. In regions with high gradients (e.g., near a surface in a turbulent flow), these terms may be very significant and the partial derivatives would be estimated using whatever (X_i, r_i) data is available.

There is also a very real possibility that measurements of different variables might share identical bias errors. This is easy to imagine for measurements of x, y, and z. Another possibility is D and S sharing an identical error source, for example if the same

density table (curve fit) is used both in data reduction in the experiment and in the simulation. In such cases, additional correlated bias terms must be included in equation (31), (32), (33), and (35).

To estimate U_{SPD} for a case in which the simulation uses previous data D_i in m instances, one would need to evaluate

$$U_{SPD}^2 = \sum_{i=1}^{m} \left(\frac{\partial S}{\partial D_i} \right)^2 \left(U_{D_i} \right)^2 \tag{37}$$

where the U_{D_i} are the uncertainties associated with the data.

3.3.2 Single CFD Code

Consideration of equation (32) shows that (1) the more uncertain the data, and/or (2) the more inaccurate the code (greater U_{SN} and U_{SPD}), the easier it is to validate a code, since the greater the uncertainties in the data and the code predictions, the greater the noise level U_V. However, if the value of U_V is greater than that designated as necessary in a research/design/development program, the required <u>level of validation</u> could not be achieved without improvement in the quality of the data, the code, or both. Also, if U_{SN} and U_{SPD} are not estimated, but $|E|$ is less than U_D, then a type of validation can be argued to have been achieved, <u>but clearly as shown by the present methodology, at an unknown level.</u>

If there is a programmatic validation requirement, denote it as U_{reqd} since validation is required at that uncertainty level or below. From a general perspective, if we consider the three variables U_V, $|E|$, and U_{reqd} there are six combinations (assuming none of the three variables are equal):

    1. $|E| < U_V < U_{reqd}$

    2. $|E| < U_{reqd} < U_V$

    3. $U_{reqd} < |E| < U_V$

    4. $U_V < |E| < U_{reqd}$ (38)

    5. $U_V < U_{reqd} < |E|$

    6. $U_{reqd} < U_V < |E|$

In cases 1, 2 and 3, $|E| < U_V$; validation is achieved at the U_V level; and the comparison error is below the noise level, so attempting to decrease the error d_{SMA} due to the modeling assumptions in the simulation is not feasible from an uncertainty standpoint.

In case 1, validation has been achieved at a level below U_{reqd}, so validation is successful from a programmatic standpoint.

In cases 4, 5 and 6, $U_V < |E|$, so the comparison error is above the noise level and using the sign and magnitude of E to estimate d_{SMA} is feasible from an uncertainty standpoint. If $U_V << |E|$, then E corresponds to d_{SMA} and the error from the modeling assumptions can be determined unambiguously. In case 4, validation is successful at the $|E|$ level from a programmatic standpoint.

A similar comparison table can be constructed using $|E_C|$, U_{V_C}, and U_{reqd}. Since E_C can be larger or smaller than E, but U_{V_C} should always be less than U_V, the results for a given corrected case are not necessarily analogous to those for the corresponding uncorrected case. That is, a variable can be validated in the corrected but not in the uncorrected case, or vice versa. However, the band $E_C \pm U_{E_C}$ should always give a smaller (therefore better) range within which the true value of E lies than the band $E - U_E$, assuming that one's confidence in using the estimate d_{SN}^* is not misplaced. Furthermore, for cases 4, 5, and 6, one can argue that E_C more likely corresponds to d_{SMA}.

In general, validation of a code's predictions of a number (N) of different variables is desired, and this means that in a particular validation effort there could be N different E, E_C, U_V, U_{V_C}, and U_{reqd} values and (perhaps) some successful and some unsuccessful validations. For each variable, a plot of the simulation prediction versus X compared with the (X_i, r_i) data points gives a traditional overview of the validation status, but the interpretation of the comparison is greatly affected by choice of the scale and the size of the symbols. A plot of $\pm U_V (\pm U_{V_C})$ and E (E_C), and U_{reqd} (if known) versus X for each variable is particularly useful in drawing conclusions, and the interpretation of the comparison is more insensitive to scale and symbol size choices.

3.3.3 Comparison of Multiple Codes and/or Models

When a validation effort involves multiple codes and/or models, the procedure discussed above -- comparison of values of E and U_V (and U_{reqd} if known) for the N variables -- should be performed for each code/model.

Since each code/model may have a different U_V, some method to compare the different codes'/models' performance for each variable in the validation is useful. The range within which (95 times out of 100) the true value of E lies is $E \pm U_E$. From equation (32), when U_{SMA} is zero then $U_V = U_E$, so for that ideal condition the maximum absolute magnitude of the 95% confidence interval is given by $|E| + U_V$. Comparison of the ($|E| + U_V$)'s for the different codes/models then shows which has the smallest range of

likely error assuming all U_{SMA}'s are zero. This allows appropriate comparisons of (low E)/(high U_V) with (high E/low U_V) codes/models.

A similar discussion holds if the corrected values are used.

3.3.4 Predictions of Trends

In some instances, the ability of a code or model to predict the trend of a variable may be the subject of a validation effort. An example would be the difference in drag for two ship configurations tested at the same Froude number. The procedure discussed above -- comparison of |E| and U_V for the drag -- should be performed for each configuration. The difference D in drag for the two configurations should then be considered as the variable that is the subject of the validation. As discussed in Coleman and Steele (1999), because of correlated bias uncertainty effects in the experimental data the magnitude of the uncertainty in D may be significantly less than the uncertainty in either of the two experimentally determined drag values. This means that the value of U_V for D may be significantly less than the U_V's for the drag values, allowing for a more stringent validation criterion for the difference than for the absolute magnitudes of the variables. Choice of the corrected or uncorrected approach should be made on a specific case-by-case basis.

3.3.5 Corrected vs. Uncorrected Simulation Results

If a validation using the corrected approach is successful at a set condition, then if one chooses to associate that validation uncertainty level with the simulation's prediction at a neighboring condition that prediction must also be corrected. That means enough runs are required at the new condition to allow estimation of the numerical errors and uncertainties. If this is not done, then the comparison error E and validation uncertainty U_V corresponding to the use of the uncorrected S and its associated (larger) U_{SN} should be the ones considered in the validation with which one wants to associate the prediction at a new condition. (Whether to and how to associate an uncertainty level at a validated condition with a prediction at a neighboring condition is very much unresolved and is justifiably the subject of much debate at this time.)

As discussed in Section 3.3.2, however, the band $E_C \pm U_{E_C}$ should always give a smaller (therefore better) range within which the true value of E lies than the band $E - U_E$, assuming that one's confidence in using the estimate δ_{SN}^* is not misplaced.

4.0 Example for RANS CFD Code

Example results of verification and validation are presented for a single CFD code and for specified objectives, geometry, conditions, and available benchmark information. The CFD code is CFDSHIP-IOWA, which is a general-purpose, multi-block, high performance computing (parallel), unsteady RANS code (Paterson et al, 1998; Wilson et al., 1998) developed for computational ship hydrodynamics. The RANS equations are solved using higher-order upwind finite differences, PISO, Baldwin-Lomax turbulence model, and exact and approximate treatments, respectively, of the kinematic and dynamic free-surface boundary conditions. The objectives are to demonstrate the usefulness of the proposed verification and validation procedures and methodology and establish the levels of verification and validation of the simulation results for an established benchmark for ship hydrodynamics CFD validation.

4.1 Geometry, Conditions, and Benchmark Data

The geometry is the Series 60 cargo/container ship. The Series 60 was used for two of the three test cases at the last international workshop on validation of ship hydrodynamics CFD codes (CFD Workshop Tokyo, 1994). The conditions for the calculations are Froude number $Fr = 0.316$, Reynolds number $Re = 4.3 \times 10^6$, and zero sinkage and trim. These are the same conditions as the experiments, except the resistance and sinkage and trim tests, as explained next. The variables selected for verification and validation are resistance C_T (integral variable) and wave profile z (point variable).

The benchmark data is provided by Toda et al. (1992), which was also the data used for the Series 60 test cases at the CFD Workshop Tokyo (1994). The data includes resistance and sinkage and trim for a range of Fr for the model free condition (i.e., free to sink and trim); and wave profiles, near-field wave pattern, and mean velocities and pressures at numerous stations from the bow to the stern and near wake, all for $Fr = (0.16, 0.316)$ and the zero sinkage and trim model fixed condition. The data also includes uncertainty estimates, which were recently confirmed/updated by Longo and Stern (1999) closely following standard procedures (Coleman and Steele, 1999).

The resistance is known to be larger for free vs. fixed models. Data for the Series 60 indicates about an 8% increase in C_T for the free vs. fixed condition over a range of Fr including Fr=0.316 (Ogiwara and Kajatani, 1994). The Toda et al. (1992) resistance values were calibrated (i.e., reduced by 8%) for effects of sinkage and trim for the present comparisons.

4.2 Computational Grids

Grid studies were conducted using four grids (m=4), which enables two separate grid studies to be performed and compared. Grid study 1 gives estimates for grid errors and uncertainties on grid 1 using the three finest grids 1-3 while grid study 2 gives estimates for grid errors and uncertainties on grid 2 using the three coarsest grids 2-4. The results for grid study 1 are given in detail and the differences for grid study 2 are also mentioned. The grids were generated using the commercial code GRIDGEN (Pointwise,

Inc.) with consideration to topology; number of points and grid refinement ratio r_G; near-wall spacing and turbulence model requirement that first point should be at $y^+<1$; bow and stern spacing; and free-surface spacing.

The topology is body-fitted, *H*-type, and single block.. The sizes of grids 1 (finest) through 4 (coarsest) are $287x78x43 = 876,211$, $201x51x31 = 317,781$, $144x36x22 = 114,048$, and $101x26x16 = 42,016$, and the grid refinement ratio $r_G = \sqrt{2}$. Clustering was used near the bow and stern in the x-direction, at the hull in the h-direction, and near the free surface in the z-direction. The y^+ values for grids 1-4 were about 0.7, 1, 1.4, and 2, respectively. About twice the number of grid points in the h-direction would be required to achieve $y^+ < 1.0$ for all four grids 1-4 (i.e., roughly 1,800,000 points on the finest grid). With grid refinement ratio $r_G = \sqrt{2}$, only grids 1 and 2 were generated. Grids 3 and 4 were obtained by removing every other point from grids 1 and 2, respectively (i.e., the grid spacing of grids 3 and 4 is twice that of grids 1 and 2, respectively). Grids 1 and 2 were generated by specifying the grid spacing at the corners and number of points along the edges of the computational blocks. The faces of the computational blocks were smoothed using an elliptic solver after which the coordinates in the interior were obtained using transfinite interpolation from the block faces. Grid 2 was generated from grid 1 by increasing the grid spacing and decreasing the number of computational cells in each coordinate direction at the corners of the blocks by a factor r_G. A comparison of the four grids at the free surface plane is shown in figure 2 along with computed wave elevation contours

4.3 Verification and Validation of Integral Variable: Resistance

Verification. Verification was performed with consideration to iterative and grid convergence studies, i.e., $d_{SN} = d_I + d_G$ and $U_{SN}^2 = U_I^2 + U_G^2$.

Iterative convergence was assessed by examining iterative history of ship forces and L2 norm of solution changes summed over all grid points. Figure 3 shows a portion of the iterative history on grid 1. The portion shown represents a computation started from a previous solution and does not reflect the total iterative history. Solution change drops four orders of magnitude from an initial value of about 10^{-2} (not shown) to a final value of 10^{-6}. The variation in C_T is about $0.07\%S_G$ over the last period of oscillation (i.e., $U_I = 0.07\%S_G$). Iterative uncertainty is estimated as half the range of the maximum and minimum values over the last two periods of oscillation (see figure 3c). Iterative histories for grids 2-4 show iterative uncertainties of about 0.02, 0.03, and $0.01\%S_G$, respectively. The level of iterative uncertainties U_I for grids 2-4 are at least two orders of magnitude less than the corresponding grid uncertainties U_G, whereas the iterative uncertainty for grid 1 is only one order of magnitude smaller than the grid error. For all four grids the iteration errors and uncertainties are assumed to be negligible in comparison to the grid errors and uncertainties for all four solutions (i.e., $d_I << d_G$ and $U_I << U_G$ such that $d_{SN} = d_G$ and $U_{SN} = U_G$).

The results from the grid convergence study for C_T are summarized in tables 2 and 3. The solutions for C_T indicate the converging condition (i) of equation (16) with

$R_G = \varepsilon_{21} / \varepsilon_{32} = 0.21$. The first-order RE estimate $d_{RE_{G_1}}$ [in equation (22)], order of accuracy p_G [in equation (23)], and correction factor C_G [in equation (24a)] are

$$d^*_{RE_{G_1}} = \left(\frac{e_{21_G}}{r_G^{p_G} - 1} \right) = \left(\frac{0.06x10^{-3}}{(\sqrt{2})^{4.4} - 1} \right) = 0.02x10^{-3} \tag{39}$$

$$p_G = \frac{\ln(e_{32_G} / e_{21_G})}{\ln(r_G)} = \frac{\ln(0.28/0.06)}{\ln(\sqrt{2})} = 4.4 \tag{40}$$

$$C_G = \frac{r_G^{p_G} - 1}{r_G^{p_{G_{est}}} - 1} = \frac{(\sqrt{2})^{4.4} - 1}{(\sqrt{2})^2 - 1} = 3.7 \tag{41}$$

where $p_{est} = p_{th} = 2$ was used in equation (41). Uncertainty and error estimates are made next both considering C_G as sufficiently less than or greater than 1 and lacking confidence and C_G as close to 1 and having confidence, as discussed in Section 3.2.3.

For $C_G = 3.7$ considered as sufficiently less than or greater than 1 and lacking confidence, U_G is estimated and not d_G

$$U_G = \left| C_G d^*_{RE_{G_1}} \right| + \left| (1 - C_G) d^*_{RE_{G_1}} \right| = 0.06x10^{-3} + 0.05x10^{-3} = 0.11x10^{-3} \tag{42}$$

U_G is 2.1% S_{G_1}.

For $C_G = 3.7$ considered close to 1 and having confidence, both and d^*_G and U_{G_C} are estimated

$$d^*_{G_1} = C_G d^*_{RE_{G_1}} = 0.06x10^{-3} \tag{43}$$

$$U_{G_C} = \left| (1 - C_G) d^*_{RE_{G_1}} \right| = 0.05x10^{-3} \tag{44}$$

The corrected solution S_C is defined with $S = S_{G_1}$

$$S_C = S_{G_1} - d^*_{G_1} = 4.99x10^{-3} \tag{45}$$

$d^*_{G_1}$ and U_{G_C} are 1.2% and 1.0% S_C, respectively. In both cases, the level of verification is relatively small <2.1% S_{G_1}.

Table 3 includes results for grid study 2, which are similar to those for grid study 1, but the values are larger by a factor of about 3, except S_C which differs by only 3%. Also shown in table 2 are the pressure C_P and frictional C_F components of C_T. C_F comprises about 70% of C_T and also displays convergence; however, C_P is convergent for the second grid study and neutrally convergent ($R_G=0$) to three significant figures for the first grid study (i.e., C_P is grid independent on the finest grid). Solution changes between grids 1 and 2 for C_P are at or below the level of iteration uncertainty ($0.1\% S_G$), so that further grid refinement is unwarranted. Apparently for this geometry, convergence of C_F with grid refinement is slower than that of C_P. The results show that the use of finer grids is problematic; since, the next largest grid with $r_G = \sqrt{2}$ would have 2.4M grid points and

20

iterative errors and grid errors would likely be of similar order of magnitude. The R_G, p_G, and C_G values are far from their asymptotic range values of $R_G=0.5$, $p_G=2$, and $C_G=1$, respectively.

Validation. Validation is performed using both the simulation prediction S and the corrected simulation prediction S_C, as summarized in table 4. First using S, the comparison error is calculated from equation (30) with $S = S_{G_1}$ as

$$E = D - S = 5.42x10^{-3} - 5.05x10^{-3} = 0.37x10^{-3} = 6.8\%D \tag{46}$$

The validation uncertainty is calculated from equation (33) as

$$U_V = \sqrt{U_{SN}^2 + U_D^2} = 0.17x10^{-3} = 3.1\%D \tag{47}$$

where $U_{SN}=U_G=1.9\%D$ and $U_D=2.5\%D$. Comparison error $|E|>U_V$ such that the simulation results are not validated. U_{SN} and U_D are of similar order such that reduction in U_V would require reduction of U_D and U_{SN}. Reduction of U_{SN} by using finer grids may be possible; however, as already mentioned, iterative errors will likely be of similar order of magnitude and will also need to be accurately estimated. E is positive, i.e., the simulation under predicts the data. The trends shown in table 2 suggest C_p too small. Presumably modeling errors such as resolution of the wave field and inclusion of effects of sinkage and trim can be addressed to reduce E and validate C_T at $U_V=3.1\%D$; however, the case for this reasoning is stronger when considering the corrected comparison error, as discussed next.

Second using S_C, the corrected comparison error is calculated from equation (34) as

$$E_C = D - S_C = 5.42x10^{-3} - 4.99x10^{-3} = 0.43x10^{-3} = 7.9\%D \tag{48}$$

The validation uncertainty is calculated from equation (35) as

$$U_{V_C} = \sqrt{U_{S_CN}^2 + U_D^2} = 0.14x10^{-3} = 2.6\%D \tag{49}$$

where $U_{S_CN} = U_{G_C} = 0.8\%D$. Here again, $|E_C|>U_{V_C}$ such that the simulation results are not validated. However, validation uncertainty U_{V_C} is relatively small and $U_{S_CN}<<U_D$ more strongly suggests than was the case for E that E_C is mostly due to modeling errors. Therefore modeling issues should/can be improved to reduce E_C and validate C_T at the reduced level $U_{V_C}=2.6\%D$ in comparison to equation (47).

The results from grid study 2 are summarized in table 5. Note that validation of the comparison error E is achieved at the level of $U_V=6.7\%D$ while validation of the corrected comparison error E_C is not.

4.4 Verification and Validation of a Point Variable: Wave Profile

Verification. Verification for the wave profile was conducted as per that described for the resistance in Section 4.3 with the distinction that a point variable is

defined over a distribution of grid points. Interpolation of the wave profile on all grids onto a common distribution is required to compute solution changes. Since calculation of the comparison error $E=D-S$ is required for validation, wave profiles on grids 1-4 are interpolated onto the distribution of the data. The same four grids were used and, here again iteration errors and uncertainties were negligible in comparison to the grid errors and uncertainties for all four solutions, i.e., $d_I << d_G$ and $U_I << U_G$ such that $d_{SN} = d_G$ and $U_{SN} = U_G$.

R_G at local maximums and minimums (i.e., $x/L = 0.1$, 0.4, and 0.65 in figure 4a) and based on L2 norm solution changes both show convergence. The spatial order of accuracy for the wave profile was computed from the L2 norm of solution changes

$$\langle p_G \rangle = \frac{\ln\left(\left\|e_{32_G}\right\|_2 / \left\|e_{21_G}\right\|_2\right)}{\ln(r_G)} = 1.3 \tag{50}$$

where $< \,\, >$ is used to denote a profile-averaged value and $\left\|e\right\|_2$ denotes the L2 norm of solution change over the N points in the region, $0 < x/L < 1$

$$\left\|e\right\|_2 = \left[\sum_{i=1}^{N} e_i^2\right]^{1/2} \tag{51}$$

Correction factor is computed from equation (24a) using order of accuracy p_G in equation (50) and $p_{G_{est}} = 2.0$

$$\langle C_G \rangle = \frac{r_G^{\langle p_G \rangle} - 1}{r_G^{p_{G_{est}}} - 1} = \frac{(\sqrt{2})^{1.3} - 1}{(\sqrt{2})^2 - 1} = 0.56 \tag{52}$$

The estimates for order of accuracy and correction factor in equations (50) and (51) were used to estimate grid error and uncertainty for the wave profile at each grid point.

For $<C_G> = 0.56$ considered as sufficiently less than or greater than 1 and lacking confidence, pointwise values for U_G are estimated and not d_G. Equation (26) is used to estimate U_G

$$U_G = \left|\langle C_G \rangle \left(\frac{e_{21_G}}{r_G^{\langle p_G \rangle} - 1}\right)\right| + \left|(1 - \langle C_G \rangle)\left(\frac{e_{21_G}}{r_G^{\langle p_G \rangle} - 1}\right)\right| \tag{53}$$

For $<C_G>=0.56$ considered close to 1 and having confidence, pointwise values for both d_G^* and U_{G_C} are estimated using equations (25) and (27)

$$d_{G_1}^* = \langle C_G \rangle \left(\frac{e_{21_G}}{r_G^{\langle p_G \rangle} - 1}\right) \tag{54}$$

$$U_G = \left|(1 - \langle C_G \rangle)\left(\frac{e_{21_G}}{r_G^{\langle p_G \rangle} - 1}\right)\right| \tag{55}$$

Equation (10) is used to calculate S_C at each grid point

$$S_C = S_{G_1} - \mathrm{d}_{G_1}^*$$ (56)

The results are summarized in table 6. The level of verification is similar to that for C_T with slightly higher values. Table 6 includes results for grid study 2, which are closer to those for grid study 1 than was the case for C_T, i.e., are only larger by a factor of 2 vs. 3 for C_T. The R_G, p_G, and C_G values are closer to and seem to be approaching the asymptotic range.

Validation. Validation of the wave profile is performed using both the simulation prediction S and the corrected simulation prediction S_C. Profile-averaged values for both definitions of the comparison error, validation uncertainty, and simulation uncertainty are given in table 7. Values are normalized with the maximum value for the wave profile $z_{max}=0.014$ and the uncertainty in the data was reported to be $3.7\% z_{max}$. For grid study 1, E is nearly validated at about 5%. The trends are similar to those for C_T, except there are smaller differences between the use of E and E_C.

The point comparison error $E=D\text{-}S$ is compared to validation uncertainty U_V in figure 4b, while error $E_C=D\text{-}S_C$ is compared to validation uncertainty U_V in figure 4d. In the latter case, the validation uncertainty U_V in figure 4d is mostly due to U_D. Much of the profile is validated. The largest errors are at the crests and trough regions, i.e., bow, shoulder, and stern waves.

The results from grid study 2 are summarized in table 8 and included in Figure 4. The results are similar to those for grid study 1, but both E and E_C and U_V and U_{V_C} are larger.

5. Conclusions and Recommendations

The verification and validation procedures and methodology presented should have applicability to a fairly broad range of CFD codes, including RANS, Navier-Stokes, Euler, boundary-element methods, and others. The concepts and definitions and associated mathematical framework are well founded. However, clearly much more work is needed for other CFD codes (such as large-eddy simulations), additional error sources, and alternative error and uncertainty estimation methods, e.g., single-grid methods and alternative strategies to account for the effects of higher-order terms in RE. Furthermore, more experience is needed through application for different codes and geometry and conditions.

Nonetheless, the verification and validation procedures and methodology are recommended for use. Use of such procedures and methodology should be helpful in guiding future developments in CFD through documentation, verification, and validation studies and in transition of CFD codes to design through establishment of credibility. Presumably, with a sufficient number of documented, verified, and validated solutions along with selected verification studies a CFD code can be accredited for a certain range of applications. The contribution of the present work is in providing procedures and methodology for the former, which hopefully will help lead to the latter.

References

AIAA, 1998, <u>Guide for the Verification and Validation of Computational Fluid Dynamics Simulations</u>, G-077-1998.

Coleman, H.W. and Steele, W.G., 1999, <u>Experimentation and Uncertainty Analysis for Engineers</u>, 2nd Edition, John Wiley & Sons, Inc., New York, NY.

Coleman, H.W. and Stern, F., 1997, "Uncertainties in CFD Code Validation," ASME J. Fluids Eng., Vol. 119, pp. 795-803. (Also see "Authors' Closure," ASME J. Fluids Eng., Vol. 120, September 1998, pp. 635-636.)

Coleman, H.W., Stern, F., Di Mascio, A., and Campana, E., 1999, "The Problem with Oscillatory Behavior in Grid Convergence Studies," technical note, submitted to ASME J. Fluids Eng., Aug. 1999.

CFD Workshop Tokyo 1994, 1994, Proceedings, Vol. 1 and 2, 1994, Ship Research Institute Ministry of Transport Ship & Ocean Foundation.

Ferziger, J.H. and Peric, M., 1996, <u>Computational Methods for Fluid Dynamics</u>, Springer-Verlag, New York.

ITTC, 1996, 21st ITTC Proceedings, "Report of the Resistance Committee", Bergen/Trondheim, Norway.

ITTC, 1999, 22nd ITTC Proceedings, "Report of the Resistance Committee," Seoul, Korea/Beijing, China.

Longo, J. and Stern, F., "Resistance, Sinkage and Trim, Wave Profile, and Nominal Wake and Uncertainty Assessment for DTMB Model 5512," Proc. 25th ATTC, Iowa City, IA, 24-25 September 1998.

Mehta, U.B., 1998, "Credible Computational Fluids Dynamics Simulations," AIAA Journal, Vol. 36, pp. 665-667.

Ogiwara, S. and Kajitani, H., 1994, "Pressure Distribution on the Hull Surface of Series 60 (CB=0.60) Model," Proceedings CFD Workshop Tokyo, Vol. 1, pp. 350-358.

Paterson, E.G., Wilson, R.V., and Stern, F., 1998, "CFDSHIP-IOWA and Steady Flow RANS Simulation of DTMB Model 5415," 1st Symposium on Marine Applications of Computational Fluid Dynamics, McLean, VA, 19-21 May.

Roache, P.J., 1998, <u>Verification and Validation in Computational Science and Engineering</u>, Hermosa publishers, Albuquerque, New Mexico.

Rood, E.P., 1996, "Validation Strategy for RANS Computational Ship Hydrodynamics," 2nd International Conference on Hydrodynamics, Hong Kong.

Shimazaki, K., Himeno, Y., and Baba, N., 1993, "An Attempt at Evaluating Numerical Errors in a 2D Navier-Stokes Solver," FED Vol. 158, Symposium on Quantification of Uncertainty in Computational Fluid Dynamics, ASME Fluids Engineering Division, Summer Meeting, Washington, D.C., pp. 19-28.

Stern, F., Paterson, E.G., and Tahara, Y., 1996, "CFDSHIP-IOWA: Computational Fluid Dynamics Method for Surface-Ship Boundary Layers and Wakes and Wave

Fields," Iowa Institute of Hydraulic Research, The University of Iowa, Iowa City, IIHR Report No. 381.

Toda, Y., Stern, F., and Longo, J., 1992, "Mean-Flow Measurements in the Boundary Layer and Wake and Wave Field of a Series 60 CB = .6 Model Ship - Part 1: Froude Numbers .16 and .316," Journal of Ship Research, Vol. 36, No. 4, pp. 360-377.

Wilson, R., Paterson, E., and Stern, F., 1998 "Unsteady RANS CFD Method for Naval Combatant in Waves," Proc. 22nd ONR Symposium on Naval Hydro, Washington, DC.

Table 1. Definitions of errors and uncertainties and verification and validation

Errors	
Present and Roache (1998)	Error d is the difference between a simulation value or an experimental value and the truth
AIAA (1998)	A recognizable deficiency in any phase or activity of modeling and simulation that is not due to lack of knowledge
Uncertainties	
Present and Roache (1998)	An uncertainty U is an estimate of an error such that the interval $\pm U$ contains the true value of d 95 times out of 100
AIAA (1998)	A potential deficiency in any phase or activity of the modeling process that is due to lack of knowledge
Verification	
Present	Verification is defined as a process for assessing numerical uncertainty U_{SN} and, when conditions permit, estimating the sign and magnitude of the numerical error d_{SN}^{*} itself and the uncertainty $U_{S_C N}$ in that error estimate.
Roache (1998)	Solving the equations right/mathematics
AIAA (1998)	The process of determining that a model implementation accurately represents the developer's conceptual description of the model and the solution to the model
Validation	
Present	Validation is defined as a process for assessing modeling uncertainty U_{SM} by using benchmark experimental data and, when conditions permit, estimating the sign and magnitude of the modeling error d_{SM} itself.
Roache (1998)	Solving the right equations/science/engineering
AIAA (1998)	The process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model

Table 2 Grid convergence study for total C_T, pressure C_P, and frictional C_F resistance ($x10^{-3}$) for Series 60.

Grid	Grid 4 101x26x16	Grid 3 144x36x22	Grid 2 201x51x31	Grid 1 287x71x43	Data
C_T	6.02	5.39	5.11	5.05	5.42
e		-10%	-5.2%	-1.2%	
C_P	1.88	1.61	1.60	1.60	$C_R = 2.00$
e		-14%	-0.6%	0.0%	
C_F	4.14	3.69	3.51	3.45	3.42
e		-11%	-4.9%	-1.7%	ITTC

% S_G.

Table 3. Verification of total resistance C_T ($x10^{-3}$) for Series 60.

Study	R_G	p_G	C_G	U_G	d_G^*	U_{G_C}	S_C
1 (grids 1-3)	0.21	4.4	3.7	2.1%	1.2%	0.9%	4.99
2 (grids 2-4)	0.44	2.3	1.3	6.7%	5.5%	1.1%	4.83

%S_G.

Table 4. Validation of total resistance for Series 60 – study 1 (grids 1-3).

	$E\%$	$U_V\%$	$U_D\%$	$U_{SN}\%$
$E=D\text{-}S$	6.8	3.1	2.5	1.9
$E_C=D\text{-}S_C$	7.9	2.6	2.5	0.8

%D.

Table 5. Validation of total resistance for Series 60 – study 2 (grids 2-4).

	$E\%$	$U_V\%$	$U_D\%$	$U_{SN}\%$
$E=D\text{-}S$	5.7	6.7	2.5	6.3
$E_C=D\text{-}S_C$	11	2.7	2.5	1.0

%D.

Table 6 Profile-averaged values from verification of wave profile for Series 60.

Study	R_G	p_G	C_G	U_G	U_{G_C}
1 (grids 1-3)	0.64	1.3	0.56	2.0%	0.9%
2 (grids 2-4)	0.68	1.1	0.47	4.1%	2.2%

$\%\zeta_{\max}$.

Table 7. Profile-averaged values from validation
of wave profile for Series 60 – study 1 (grids 1-3).

	$E\%$	$U_V\%$	$U_D\%$	$U_{SN}\%$
$E=D\text{-}S$	5.2	4.2	3.7	2.0
$E_C=D\text{-}S_C$	5.6	3.8	3.7	0.9

$\%\zeta_{\max}$.

Table 8. Profile-averaged values from validation
of wave profile for Series 60 – study 2 (grids 2-4).

	$E\%$	$U_V\%$	$U_D\%$	$U_{SN}\%$
$E=D\text{-}S$	5.6	5.5	3.7	4.1
$E_C=D\text{-}S_C$	6.6	4.3	3.7	2.2

$\%\zeta_{max}$.
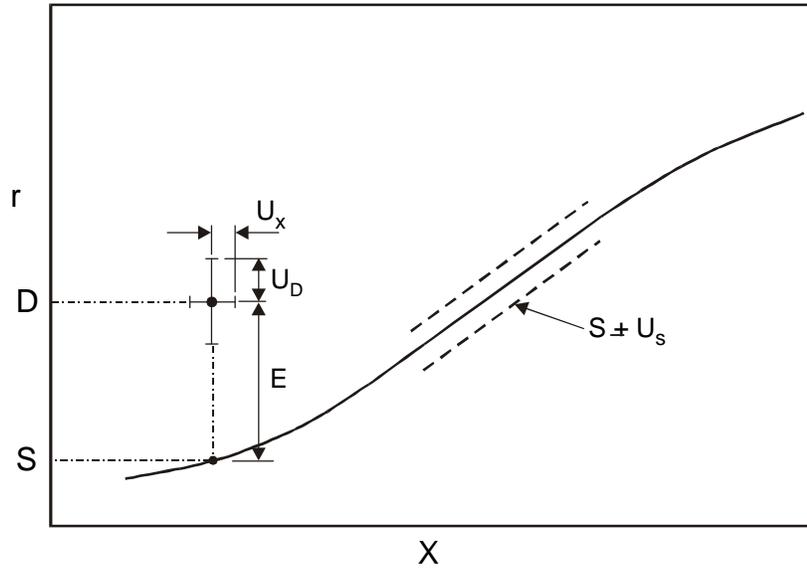
Figure 1 Definition of comparison error.
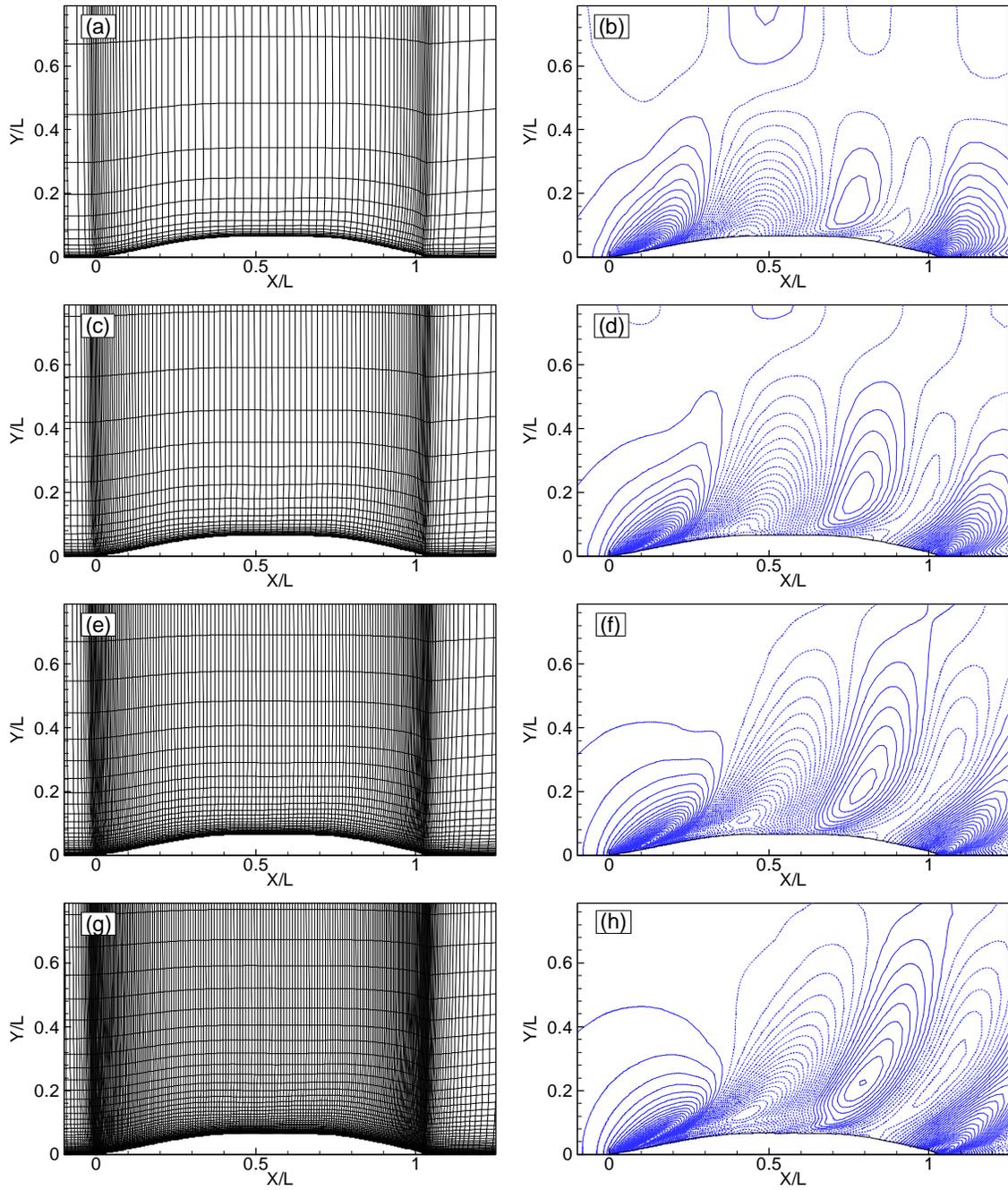
Figure 2. Grids and wave contours from verification and validation studies for Series 60: (a) and (b) coarsest - grid 4; (c) and (d) grid 3; (e) and (f) grid 2; and (g) and (h) finest - grid 1.

30
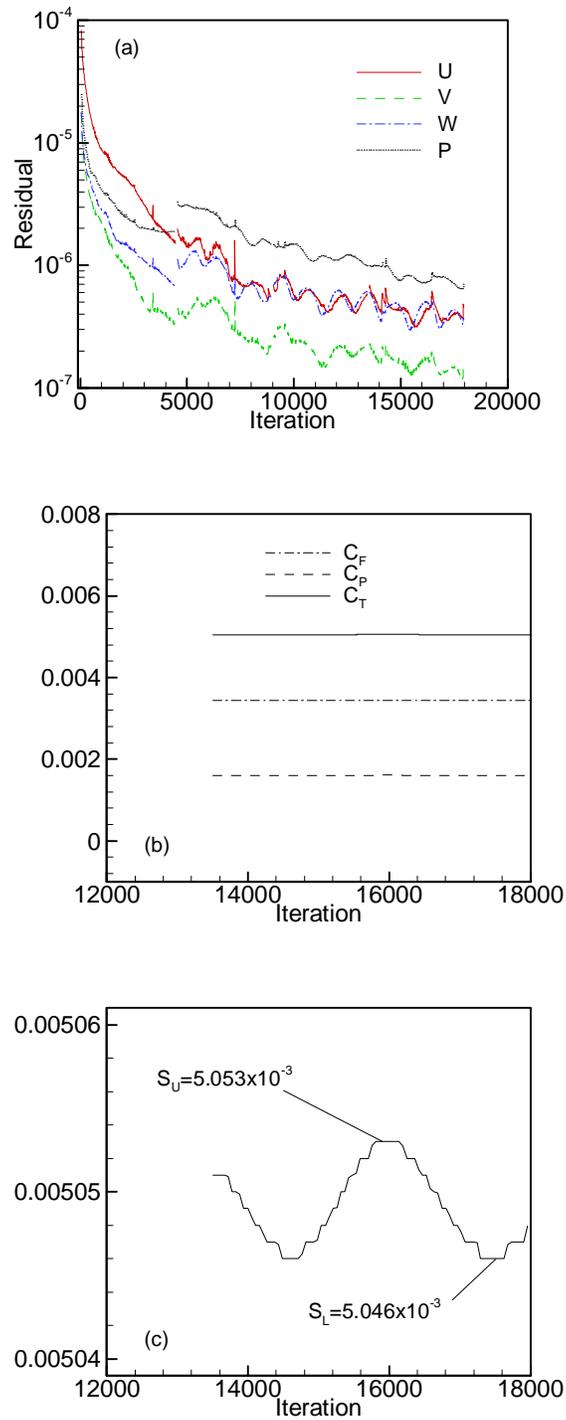
Figure 3. Iteration history for Series 60 on grid 1: (a) solution change, (b) ship forces - C_F, C_P, and C_T and (c) magnified view of total resistance C_T over last two periods of oscillation.

31
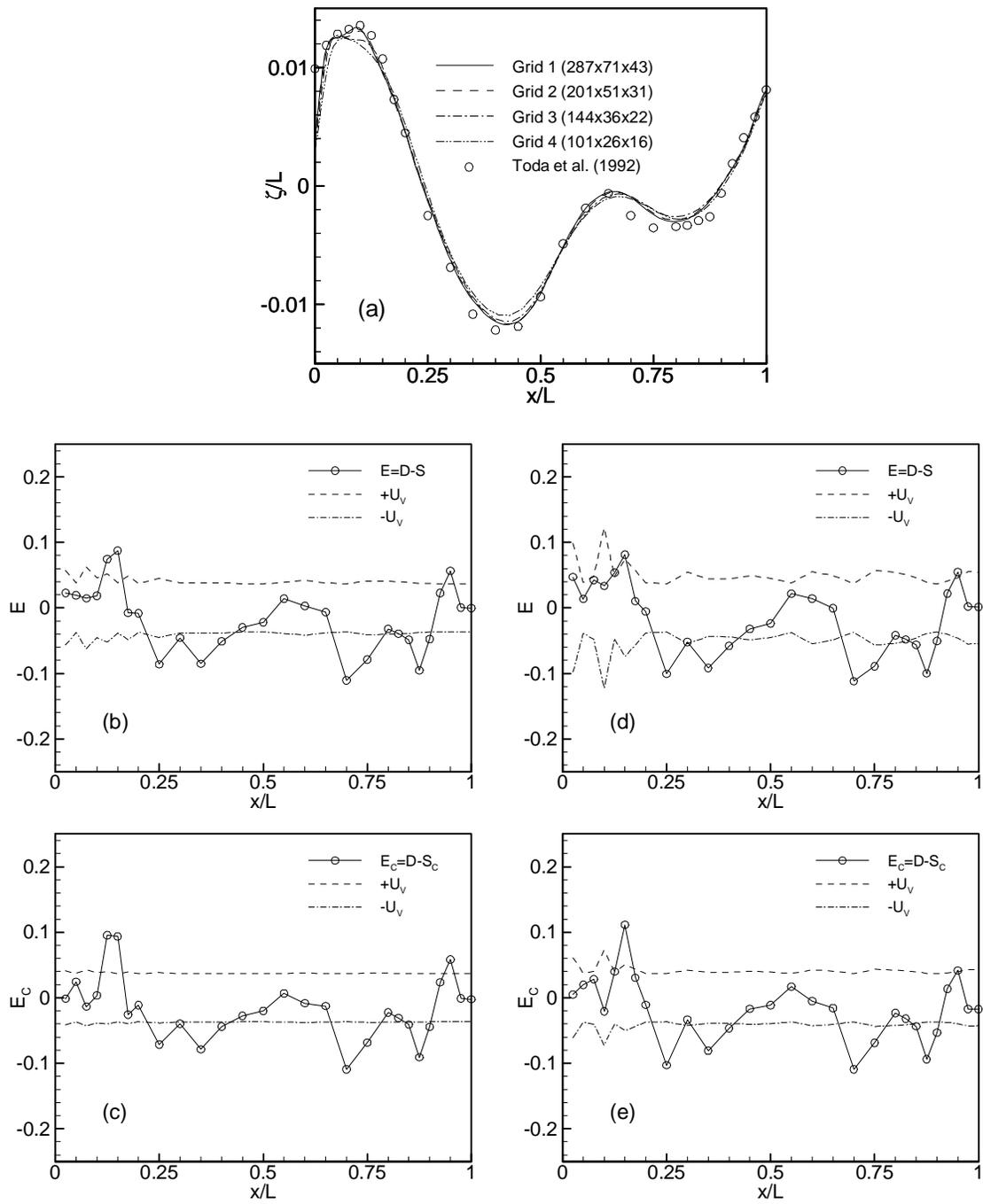
Figure 4. Wave profile for Series 60: (a) grid study; (b) and (c) validation using grids 1-3; and (d) and (e) validation using grids 2-4.

Appendix A. Derivation of Simulation Error Equation

There are three engineering approaches to solving fluid mechanics problems: analytical, experimental, and simulation. Fluid mechanics problems are governed by the laws of physics, which are formulated for unsteady flow as initial and boundary value problems (IBVP) and for steady flow problems as boundary value problems (BVP).

The IBVP is defined by a continuous partial differential equation (PDE) operator L_T with specified initial (IC) and boundary (BC) conditions

$$L_T(T) = 0$$
$$IC : T(x, t = 0) = G_T(x)$$
$$BC : T(x_B, t) = H_T(t)$$

(A.1)

x is the spatial coordinate(s) and may be a vector, the functions G_T and H_T are the IC (at $t=0$) and BC (at $x = x_B$), respectively, t is time, and T is the true or exact solution. By definition, equation (A.1) contains no modeling or numerical errors.

The experimental approach does not solve equation (A.1), but instead uses experimental measurement systems to determine T. This process results in bias and precision errors that lead to an uncertainty U_D in the experimental measurement D.

Analytical and simulation approaches formulate the IBVP by selection of the PDE, IC, and BC to model the physical phenomena

$$L_M(M) = 0$$
$$IC : M(x, t = 0) = G_M(x)$$
$$BC : M(x_B, t) = H_M(t)$$

(A.2)

with similar definitions as per equation (A.1); however, L_M, G_M, H_M, and x_B all may contain modeling assumptions such that $M \neq T$.

Assumptions are made in modeling geometry, turbulence, non-Newtonian fluids, combustion, compressibility, two-fluid and rarified gas flows, etc. An IBVP for the modeling error $\mathrm{d}_{SM} = M - T$ (i.e., the difference between the model and true values) can be obtained by subtracting equation (A.1) and (A.2), then subtracting $L_M(T)$ from both sides of that result, and lastly assuming that the operator L_M is linear

$$L_M(M - T) = L_M(\mathrm{d}_{SM}) = \Gamma_M = -L_M(T)$$
$$IC : \mathrm{d}_{SM}(x, t = 0) = G_M(x) - G_T(x)$$
$$BC : \mathrm{d}_{SM}(x_B, t) = H_M(t) - H_T(t)$$

(A.3)

The assumption that L_M is linear is a major limitation since most fluid mechanics problems of interest are governed by non-linear operators. However, a linear analysis (e.g., stability analysis of explicit methods, modified equation, convergence rates for iterative methods, etc.) is often used successfully to make the problem tractable and to provide insight into the problem of interest. Equation (A.3) shows that the modeling error d_{SM} is governed by

the same operator as the solution *M*, but with a source term and differences in the IC and BC due to errors in modeling the true operator and IC and BC functions.

The analytical approach solves equation (A.2) exactly and is thus limited to simple fluid mechanics problems. The simulation approach solves equation (A.2) approximately using numerical methods and thus introduces numerical errors. The continuous IBVP is reduced to a discrete IBVP, i.e., algebraic equations using spatial and temporal discretization techniques such as finite difference, volume, and element methods resulting in numerical errors due to spatial Δx, temporal Δt, and other step sizes Δx_j (i.e., the numerical error is zero when the step sizes are zero). The discrete IBVP is defined by a discrete operator $\mathcal{L}_N$ with discrete IC and BC

$$\mathcal{L}_N(S) = \Gamma_I$$
$$IC: S(x, t = 0) = \mathcal{G}_N(x) \tag{A.4}$$
$$BC: S(x_B, t) = \mathcal{H}_N(t)$$

where the source term G_I is the residual imbalance of the algebraic equations due to the use of implicit methods. If explicit methods are used, iterative errors do not exist and $G_I = 0$. Equation (A.4) is solved on a computer through a set of programming instructions (i.e., a CFD computer code) to provide the simulation prediction *S*. Program execution requires specification of various input parameters, including step size distributions.

Numerical errors can be defined and evaluated by transforming the discrete IBVP back to a continuous IBVP. This is accomplished by representing *S* as a generalized Taylor series about a numerical benchmark S_C (solution with zero step sizes) in terms of step sizes Δx_j

$$S = S_C + \sum_{j=1}^{J} \sum_{i=1}^{\infty} \frac{(\Delta x_j)^i}{i!} \frac{\partial^i S}{\partial x_j^i} \tag{A.5}$$

where *j=1, J* is used to represent various step sizes introduced in discretization of the continuous PDE, IC, and BC (spatial $\Delta x_G = \Delta x$, temporal $\Delta x_T = \Delta t$, and other Δx_j). Substituting expansion (A.5) into equation (A.4) and rearranging gives the modified equation that is actually solved when discretization techniques are applied to an IBVP (Anderson et al., 1984)

$$L_{Modified}(S) = L_M(S) = \Gamma_N$$
$$IC: S(x, t = 0) = G_{Modified}(x) \tag{A.6}$$
$$BC: S(x_B, t) = H_{Modified}(t)$$

where the source term is given by

$$\Gamma_N = \Gamma_I + \sum_{j=1}^{J} \Gamma_j \tag{A.7}$$

The summation term in equation (A.7) represents the truncation errors due to differences between the continuous and discrete PDE. Spatial and temporal truncation error terms for

34

typical finite difference and volume methods are in the form of a power series expansion in step sizes $\mathrm{D}x_j$

$$\Gamma_j = \sum_{i=1}^{\infty} (\Delta x_j)^{p_j^{(i)}} \mathrm{a}_j^{(i)} \tag{A.8}$$

where the superscript (i) is used to indicate variables in the i_{th} term of the expansion, $\mathrm{a}_j^{(i)} = f(\dfrac{\partial^l S}{\partial x_j^l})$ contains solution derivatives with respect the x_j and are independent of the step sizes Δx_j, and $p_j^{(i)}$ is the rate of reduction of the truncation error terms with refinement of Δx_j (i.e., order of accuracy). The modified equation (A.6) recovers the modeled operator L_M; however, it operates on the simulation prediction S instead of the exact solution to the modeled equations M. Thus, the source term G_N causes the simulation prediction S to differ from the exact solution to the modeled equation M.

Subtracting equations (A.2) and (A.6) gives the IBVP that governs the simulation numerical error $\mathrm{d}_{SN} = S - M$ (i.e., the difference between the simulation and modeled values) (Ferziger 1993; Roache 1998)

$$L_M(S - M) = L_M(\mathrm{d}_{SN}) = \Gamma_N = \Gamma_I + \sum_{j=1}^{J} \Gamma_j$$
$$IC : \mathrm{d}_{SN}(x, t = 0) = G_{Modified}(x) - G_M(x) \tag{A.9}$$
$$BC : \mathrm{d}_{SN}(x_B, t) = H_{Modified}(t) - H_M(t)$$

Thus, the iterative and truncation error terms also act as source terms for numerical errors d_{SN} in the solution S. If there are no iterative errors, the source term G_N (and thus d_{SN}) is zero when either the truncation error is zero (e.g., spectral methods) or step size Δx_j is zero. If the exact form of the truncation error terms [equation (A.8)] for a discretization technique is known, equation (A.9) can be solved numerically to give d_{SN}. Such methods can be classified as single step-size error estimation methods (e.g., Shimazaki et al., 1993).

Rewriting the PDE for the numerical error in equation (A.9) with the source term G_N expanded gives

$$L_M(\mathrm{d}_{SN}) = \Gamma_I + \sum_{j=1}^{J} \sum_{i=1}^{\infty} (\Delta x_j)^{p_j^{(i)}} \mathrm{a}_j^{(i)} \tag{A.10}$$

Since the term $(\Delta x_j)^{p_j^{(i)}}$ is independent of the continuous operator L_M, the numerical error d_{SN} [i.e., the solution to equation (A.10)] reduces at the same rate as the source term G_N in equation (A.10) so that the solution is of the form

$$\mathrm{d}_{SN} = \mathrm{d}_I + \sum_{j=1}^{J} \mathrm{d}_j \tag{A.11}$$

where

$$d_j = \sum_{i=1}^{\infty} (\Delta x_j)^{p_j^{(i)}} g_j^{(i)} \tag{A.12}$$

and $g_j^{(i)} = f(\dfrac{\partial^l S}{\partial x_j^l})$ is the "grid" function which contains continuous solution derivatives.

The form of equation (A.11) can be verified by substitution of equation (A.11) with (A.12) into (A.10), which gives

$$L_M[d_I + \sum_{j=1}^{J} \sum_{i=1}^{\infty} (\Delta x_j)^{p_j^{(i)}} g_j^{(i)}] = L_M(d_I) + \sum_{j=1}^{J} \sum_{i=1}^{\infty} L_M[(\Delta x_j)^{p_j^{(i)}} g_j^{(i)}] \tag{A.13}$$

If step size Δx_j and order of accuracy p_j are constant and independent of L_M, the last term in equation (A.13) can be rewritten

$$L_M[d_I + \sum_{j=1}^{J} \sum_{i=1}^{\infty} (\Delta x_j)^{p_j^{(i)}} g_j^{(i)}] = L_M(d_I) + \sum_{j=1}^{J} \sum_{i=1}^{\infty} (\Delta x_j)^{p_j^{(i)}} L_M[g_j^{(i)}] \tag{A.14}$$

Comparison with equation (A.10) gives the following equations

$$L_M(d_I) = \Gamma_I \tag{A.15}$$

$$L_M[g_j^{(i)}] = a_j^{(i)} \tag{A.16}$$

The solution to equation (A.16) for $g_j^{(i)}$ is independent of step size Δx_j and order of accuracy p_j. Thus, the form of the numerical error d_{SN} assumed in equation (A.11) involve products of $(\Delta x_j)^{p_j^{(i)}}$ and functions that are independent of step size and order of accuracy. This verifies that the numerical error d_{SN} reduces as $(\Delta x_j)^{p_j^{(i)}}$.

Finally, the IBVP that governs the simulation error d_S is obtained by adding equations (A.3) and (A.9)

$$\begin{aligned} L_M(S-T) = L_M(d_S) &= \Gamma_N + \Gamma_M \\ IC: d_S(x,0) &= G_{Modified}(x) - G_T(x) \\ BC: d_S(x_B,t) &= H_{Modified}(t) - H_T(t) \end{aligned} \tag{A.17}$$

where simulation error is defined as

$$d_S = S - T = d_{SN} + d_{SM} \tag{A.18}$$

Equation (A.18) provides the desired expression for the simulation error in terms of the simulation modeling and numerical errors. It shows that the simulation modeling and numerical errors are additive subject to the assumption that L_M is a linear operator.

Appendix B. Generalized Richardson Extrapolation

For the converging condition (i) in equation (16), generalized RE is used to estimate U_k or d_k^* and U_{k_C}. RE is generalized for J input parameters and for use of correction factors to account for the effects of higher-order terms and defining and estimating errors and uncertainties, as summarized in Section 3.2.3. This appendix provides a detailed description.

Generalized RE begins with equation (14). The error terms on the right-hand-side of equation (14) are of known form (i.e., power series expansion in Δx_k) based on analysis of the modified (A.6) and numerical error (A.9) equations, as shown in Appendix A equation (A.12), which is written below as a finite sum (i.e., error estimate) and for the *kth* parameter and *mth* solution

$$\mathsf{d}_{k_m}^* = \sum_{i=1}^{n} (\Delta x_{k_m})^{p_k^{(i)}} g_k^{(i)} \tag{B.1}$$

n = number of terms retained in the power series, powers $p_k^{(i)}$ correspond to order of accuracy (for the *ith* term), and $g_k^{(i)}$ are referred to as "grid" functions which are a function of various orders and combinations of derivatives of S with respect to x_k. Substituting equation (20) into equation (14) results in

$$\hat{S}_{k_m} = S_C + \sum_{i=1}^{n} (\Delta x_{k_m})^{p_k^{(i)}} g_k^{(i)} + \sum_{j=1, j \neq k}^{J} \mathsf{d}_{j_m}^* \tag{B.2}$$

Subtraction of multiple solutions where input parameter Δx_k is uniformly refined eliminates the $\mathsf{d}_{j_m}^*$ terms in equation (B.2) since $\mathsf{d}_{j_m}^*$ is independent of Δx_k and provides equations for S_C, $p_k^{(i)}$, and $g_k^{(i)}$. This assumes $p_k^{(i)}$ and $g_k^{(i)}$ are also independent of Δx_k. Since each term (i) contains 2 unknowns, $m=2n+1$ solutions are required to estimate the numerical benchmark S_C and the first n terms in the expansion in equation (B.2) (i.e., for $n=1$, $m=3$ and for $n=2$, $m=5$, etc). The accuracy of the estimates depends on how many terms are retained in equation (B.1), the magnitude (importance) of the higher-order terms, and the validity of the assumption that $p_k^{(i)}$ and $g_k^{(i)}$ are independent of Δx_k. For sufficiently small Δx_k, the solutions are in the asymptotic range such that higher–order terms are negligible and the assumption that $p_k^{(i)}$ and $g_k^{(i)}$ are independent of Δx_k is valid. However, achieving the asymptotic range for practical geometry and conditions is usually not possible and $m>3$ is undesirable from a resources point of view; therefore, methods are needed to account for effects of higher-order terms for practical application of RE. Additionally, methods may be needed to account for possible dependence of $p_k^{(i)}$ and $g_k^{(i)}$ on Δx_k, although not addressed herein. Usually d_k^* is estimated for the finest value of the input parameter, i.e., $\mathsf{d}_k^* = \mathsf{d}_{k_1}^*$ corresponding to the finest solution S_{k_1}. RE can be classified as a multiple step-size error estimation method.

If only the leading term ($n=1$) in equation (B.1) is estimated, three solutions are required and can be written from equation (B.2)

$$\hat{S}_{k_1} = S_C + \left(\Delta x_{k_1}\right)^{p_k^{(1)}} g_k^{(1)} + \sum_{j=1,i\neq k}^{J} \mathsf{d}_{j_1}^{*} \tag{B.3}$$

$$\hat{S}_{k_2} = S_C + \left(r_k \Delta x_{k_1}\right)^{p_k^{(1)}} g_k^{(1)} + \sum_{j=1,i\neq k}^{J} \mathsf{d}_{j_1}^{*} \tag{B.4}$$

$$\hat{S}_{k_3} = S_C + \left(r_k^2 \Delta x_{k_1}\right)^{p_k^{(1)}} g_k^{(1)} + \sum_{j=1,i\neq k}^{J} \mathsf{d}_{j_1}^{*} \tag{B.5}$$

Equations (B.3)-(B.5) provide three equations for the three unknowns (S_C, $p_k^{(1)}$, and $g_k^{(1)}$). The order of accuracy and "grid" function are found by computing the solution changes $\mathsf{e}_{21_k} = \hat{S}_{k_2} - \hat{S}_{k_1}$ and $\mathsf{e}_{32_k} = \hat{S}_{k_3} - \hat{S}_{k_2}$ from equations (B.3) - (B.5)

$$\mathsf{e}_{21_k} = \hat{S}_{k_2} - \hat{S}_{k_1} = \left(\Delta x_{k_1}\right)^{p_k^{(1)}} g_k^{(1)} (r_k^{p_k^{(1)}} - 1) \tag{B.6}$$

$$\mathsf{e}_{32_k} = \hat{S}_{k_3} - \hat{S}_{k_2} = r_k^{p_k^{(1)}} \left(\Delta x_{k_1}\right)^{p_k^{(1)}} g_k^{(1)} (r_k^{p_k^{(1)}} - 1) \tag{B.7}$$

The estimate of the error $\mathsf{d}_{k_1} \approx \mathsf{d}_{RE_1}^{(1)} = \left(\Delta x_{k_1}\right)^{p_k^{(1)}} g_k^{(1)}$ is obtained from equation (B.6)

$$\mathsf{d}_{RE_1}^{*(1)} = \frac{\mathsf{e}_{21_k}}{(r_k^{p_k^{(1)}} - 1)} \tag{B.8}$$

where $\mathsf{d}_{RE_1}^{*(1)}$ is an estimate of the first term of the expansion in equation (B.1) using RE. The order of accuracy $p_k^{(1)}$ is obtained by eliminating the term $\left(\Delta x_{k_1}\right)^{p_k^{(1)}} g_k^{(1)} (r_k^{p_k^{(1)}} - 1)$ from equation (B.6) and (B.7)

$$p_k^{(1)} = \frac{\ln(\mathsf{e}_{32_k} / \mathsf{e}_{21_k})}{\ln(r_k)} \tag{B.9}$$

If order of accuracy is assumed known (e.g., from the modified equation or from grid refinement tests for simple geometry using similar grid expansion) only two solutions are required to obtain an estimate of the leading term in the power series expansion equation (B.1). However, a minimum of three solutions is required to establish convergence with refinement of input parameter.

An estimate using the first two terms ($n=2$) in equation (B.1) can be obtained from five solutions

$$\hat{S}_{k_1} = S_C + \left(\Delta x_{k_1}\right)^{p_k^{(1)}} g_k^{(1)} + \left(\Delta x_{k_1}\right)^{p_k^{(2)}} g_k^{(2)} + \sum_{j=1,i\neq k}^{J} \mathsf{d}_{j_1}^{*} \tag{B.10}$$

$$\hat{S}_{k_2} = S_C + \left(r_k \Delta x_{k_1}\right)^{p_k^{(1)}} g_k^{(1)} + \left(r_k \Delta x_{k_1}\right)^{p_k^{(2)}} g_k^{(2)} + \sum_{j=1,i\neq k}^{J} d_{j_1}^* \qquad (B.11)$$

$$\hat{S}_{k_3} = S_C + \left(r_k^2 \Delta x_{k_1}\right)^{p_k^{(1)}} g_k^{(1)} + \left(r_k^2 \Delta x_{k_1}\right)^{p_k^{(2)}} g_k^{(2)} + \sum_{j=1,i\neq k}^{J} d_{j_1}^* \qquad (B.12)$$

$$\hat{S}_{k_4} = S_C + \left(r_k^3 \Delta x_{k_1}\right)^{p_k^{(1)}} g_k^{(1)} + \left(r_k^3 \Delta x_{k_1}\right)^{p_k^{(2)}} g_k^{(2)} + \sum_{j=1,i\neq k}^{J} d_{j_1}^* \qquad (B.13)$$

$$\hat{S}_{k_5} = S_C + \left(r_k^4 \Delta x_{k_1}\right)^{p_k^{(1)}} g_k^{(1)} + \left(r_k^4 \Delta x_{k_1}\right)^{p_k^{(2)}} g_k^{(2)} + \sum_{j=1,i\neq k}^{J} d_{j_1}^* \qquad (B.14)$$

The orders of accuracy and the grid functions are obtained by computing the four solution changes $e_{21_k} = \hat{S}_{k_2} - \hat{S}_{k_1}$, $e_{32_k} = \hat{S}_{k_3} - \hat{S}_{k_2}$, $e_{43_k} = \hat{S}_{k_4} - \hat{S}_{k_3}$, and $e_{54_k} = \hat{S}_{k_5} - \hat{S}_{k_4}$ which gives four equations for the four unknowns, $p_k^{(1)}$, $p_k^{(2)}$, $g_k^{(1)}$ and $g_k^{(2)}$. Upon solution, the four unknowns are used to give an estimate of the first and second terms in equation (B.1)

$$d_{RE_{k_1}}^{*(2)} = \frac{r_k^{p_k^{(2)}} e_{21_k} - e_{32_k}}{(r_k^{p_k^{(2)}} - r_k^{p_k^{(1)}})(r_k^{p_k^{(1)}} - 1)} - \frac{r_k^{p_k^{(1)}} e_{21_k} - e_{32_k}}{(r_k^{p_k^{(2)}} - r_k^{p_k^{(1)}})(r_k^{p_k^{(2)}} - 1)} \qquad (B.15)$$

where $d_{RE_1}^{*(2)}$ is an estimate of the first two terms of the expansion in equation (B.1) using RE. The orders of accuracy of the first and second terms in the expansion $p_k^{(1)}$ and $p_k^{(2)}$ are given by

$$p_k^{(1)} = \frac{\ln[a_k - b_k]}{\ln(r_k)}$$

$$p_k^{(2)} = \frac{\ln[a_k + b_k]}{\ln(r_k)} \qquad (B.16)$$

where

$$a_k = a_k / b_k$$
$$b_k = -\sqrt{c_k} / b_k$$
$$a_k = e_{21_k} e_{54_k} - e_{32_k} e_{43_k}$$
$$b_k = 2\left(e_{21_k} e_{43_k} - e_{32_k}^2\right)$$
$$c_k = -3 e_{32_k}^2 e_{43_k}^2 + 4 e_{21_k} e_{43_k}^3 + 4 e_{32_k}^3 e_{54_k}$$
$$- 6 e_{21_k} e_{32_k} e_{43_k} e_{54_k} + e_{21_k}^2 e_{54_k}^2$$

Two conditions are required to obtain estimates of the orders of accuracy from equation (B.16): (i) $a_k \pm b_k > 1$; and (ii) $c_k \geq 0$. Condition (i) is be satisfied if the solutions are monotonically convergence while condition (ii) is satisfied if the solutions are sufficiently close to the asymptotic range. If the orders of accuracy are assumed known, only three solutions are required to estimate the first two terms in the power series expansion using equation (B.15).

Appendix C. Analytical Benchmarks

Analytical benchmarks can be defined as the truth and are useful in development and confirmation of verification procedures and methodology and in code development, but can not be used for validation and are restricted to simple equations. Results were obtained for two analytical benchmarks: one-dimensional (1D) wave and two-dimensional (2D) Laplace equations. The results for the 2D Laplace equation were qualitatively similar to those for the 1D wave equation, which are presented in this appendix.

The IBVP and solutions for the true, model, and analytical benchmarks are equivalent such that the modeling error is zero and the only simulation error is the simulation numerical error. Simulation results (for the 1D wave and 2D Laplace equations) are compared to analytical benchmark solutions to determine the exact simulation numerical error and evaluate both single and multiple step-size error estimation methods. In the latter case, generalized RE is used and the role of higher-order terms in the power series expansion of the simulation numerical error [equation (B.1)] is assessed by comparing estimates of the leading term using three grids to those for the first two terms using five grids. Correction factors are derived to account for the effects of the higher-order terms and to define the uncertainty in the error estimate.

C.1 Verification Using Analytical Benchmarks

By definition, the IBVP for the true, model, and analytical benchmark solutions are equivalent

$$L_T(T) = L_M(M) = L_A(A) = 0$$

$$IC : T(x, t = 0) = M(x, t = 0) = A(x, t = 0) = G_T(x) = G_M(x) = G_A(x) \tag{C.1}$$

$$BC : T(x_B, t) = M(x_B, t) = A(x_B, t) = H_T(t) = H_M(t) = H_A(t)$$

Therefore,

$$T = M = A \tag{C.2}$$

and

$$\mathrm{d}_{SM} = 0 \tag{C.3}$$

The simulation error and uncertainty are given by

$$\mathrm{d}_S = S - A = \mathrm{d}_{SN} \tag{C.4}$$

$$U_S^2 = U_{SN}^2 \tag{C.5}$$

and the corrected simulation error and uncertainty are given by

$$\mathrm{d}_{S_C} = S_C - A = \mathrm{e}_{SN} \tag{C.6}$$

$$U_{S_C}^2 = U_{S_C N}^2 \tag{C.7}$$

Simulations are verified if

$$|E| = |A - S| < U_{SN} \tag{C.8}$$

and corrected simulations are verified if

$$|E_C| = |A - S_C| < U_{S_C N} \tag{C.9}$$

C.2 IBVP and Analytical and Numerical Solutions for 1D Wave Equation

The 1D wave equation is called a model equation, as it models the behavior of more complicated (nonlinear) PDE. A simplified form is the first-order linear convection equation with IBVP

$$L_A(A) = \frac{\partial A}{\partial t} + c\frac{\partial A}{\partial x} = 0$$

$$IC: A(x,0) = A_0 \exp\left[-\frac{(x)^2}{B}\right] \tag{C.10}$$

$$BC: A(-\infty,t) = 0$$

The initial condition is prescribed by a Gaussian function centered at $x = 0.0$ with $A_O = 1$ and $B = 0.005$, the boundary condition far upstream is zero, and c is the wave speed, which is set to unity. The computational domain is defined as $-1 \le x \le 2$.

The exact solution to equation (C.10) is

$$A(x,t) = A_0 \exp\left[-\frac{(x-ct)^2}{B}\right] \tag{C.11}$$

Figure C.1 shows the initial condition and the exact solution at $t = 1$.

Two discretization techniques are studied: (i) a first-order (Euler) explicit method; and (ii) a second-order implicit method.

For the Euler explicit method, the discrete operator in equation (A.4) is given by

$$\mathcal{L}_N(S) = \frac{S_i^{n+1} - S_i^n}{\Delta x_T} + c\frac{S_i^n - S_{i-1}^n}{\Delta x_G} = 0 \tag{C.12}$$

where $n+1$ and n denote the new and current time levels, respectively. The modified equation (A.6) and simulation numerical error equation (A.9) are given by

$$L_M(S) = \frac{\partial S}{\partial t} + c\frac{\partial S}{\partial x} = \Gamma_N \tag{C.13}$$

$$L_M(d_{SN}) = \frac{\partial d_{SN}}{\partial t} + c\frac{\partial d_{SN}}{\partial x} = \Gamma_N \tag{C.14}$$

with source terms [equation (A.8)]

$$\Gamma_N = \Delta x_G\left(\frac{cS_{xx}}{2}\right) + \Delta x_T\left(-\frac{S_{tt}}{2}\right) + O[(\Delta x_G)^2, (\Delta x_T)^2] \tag{C.15}$$

41

For equation (C.13), the IC and BC are the same as for equation (C.10). For equation (C.14), the IC and BC are given by $d_{SN}(x,0)=0$ and $d_{SN}(-\infty,t)=0$, respectively.

Similarly for the second-order implicit method,

$$\mathcal{L}_N(S) = \frac{3S_i^{n+1} - 4S_i^n + S_i^{n-1}}{2\Delta x_T} + c\frac{S_{i+1}^{n+1} - S_{i-1}^{n+1}}{2\Delta x_G} = 0 \tag{C.16}$$

$$L_M(S) = \frac{\partial S}{\partial t} + c\frac{\partial S}{\partial x} = \Gamma_N \tag{C.17}$$

$$L_M(d_{SN}) = \frac{\partial d_{SN}}{\partial t} + c\frac{\partial d_{SN}}{\partial x} = \Gamma_N \tag{C.18}$$

$$\Gamma_N = (\Delta x_G)^2\left(\frac{-cS_{xxx}}{6}\right) + (\Delta x_T)^2\left(\frac{S_{ttt}}{3}\right) + O[(\Delta x_G)^4,(\Delta x_T)^3] \tag{C.19}$$

For equations (C.17) and (C.18), the IC and BC are the same as for equations (C.13) and (C.14), respectively.

The form of the solution to equation (C.14) and (C.17) was given in Appendix A by equation (A.11) and (A.12), as a power series expansion in Dx_j. For the Euler explicit method, the form is given by

$$d_{SN}^* = d_G + d_T = \Delta x_G g_G^{(1)} + \Delta x_T g_T^{(1)} + O[(\Delta x_G)^2,(\Delta x_T)^2] \tag{C.20}$$

For the second-order implicit method, the form is given by

$$d_{SN}^* = d_G + d_T = (\Delta x_G)^2 g_G^{(1)} + (\Delta x_T)^2 g_T^{(1)} + O[(\Delta x_G)^4,(\Delta x_T)^4] \tag{C.21}$$

Results were obtained for the numerical solution of equation (C.12) and C.16) using ten grids (Table C.1) and with two values of $CFL = c\Delta t / \Delta x = (0.1, 0.5)$. The solutions were monotonically convergent for all ten grids and both CFL based on the convergence ratio R [equation (15)] defined with the ratio of the L2 norm of solution changes $\left\|e_{32_{SN}}\right\|_2$ and $\left\|e_{21_{SN}}\right\|_2$. Figure C.2 compares simulation S to analytical benchmark A for t=1 (along with single and multiple step size error estimates to be discussed later). The inherent deficiencies of the two methods are apparent. The first-order method displays dissipation errors due to even simulation derivatives in equation (C.15), which reduce at a first-order rate, whereas the second-order implicit method displays dispersion errors due to odd simulation derivatives in equation (C.19), which reduce at a second-order rate.

C.3 Single Step Size Error Estimation Method

Single step size error estimation methods are based on solution of the IBVP for the simulation numerical error d_{SN}, as given by equation (A.9). This provides an estimate of the numerical error for a single step size. Two step sizes are required to evaluate convergence with respect to input parameter. Such methods require fewer solutions than multiple step-size error estimation methods; however, there are several obstacles for practical problems. Derivation of the modified equation (A.6) is necessary in order to

define the truncation error terms [equation (A.8)], which are the source terms for equation (A.9). This may be difficult or not possible depending on the complexity and type of discretization technique used. The coefficients in the source term are functions of higher-order solution derivatives. Higher-order discretization techniques with associated increased numerical instabilities must be used to discretize the numerical error equation (A.9) than those in used in the original IBVP equation (A.4) such that the truncation error terms for the discrete form of equation (A.9) are higher order than those for equation (A.4). Also, additional programming, memory, and computer time are required to include solution of the simulation numerical error equation.

Results were obtained for the numerical solution of the numerical error equation (C.14) and (C.18) for grids 6-10 of Table C.1 and with CFL=0.1. Fourth-order spatial and third-order time discretization techniques were used. Figure C.2 compares the exact comparison error $E=A\text{-}S$ to the single step size error estimate for d_{SN}. The results show that single step size error estimates are accurate even for the first-order method and coarsest grid.

C.4 Multiple Step Size Error Estimation Method

Multiple step size error estimation methods are based on generalized RE, as described in Appendix B. The total true numerical error (i.e., grid size and time step) can be computed since the exact solution is known for the analytical benchmark. However, the exact grid size or time step error cannot be computed separately. As such, a combined grid and time step study was conducted with CFL=0.5 for all ten grids to directly compare the true error to estimates from RE. For the combined grid and time step study, overall order of accuracy (i.e., spatial *and* temporal) is estimated and the subscript *SN* is used to denote an estimate of total simulation error.

The role of higher-order terms in the power series expansion of the simulation numerical error is assessed by comparing estimates of the leading term using three grids to those for the first two terms using five grids. To avoid problems associated with pointwise calculation of order of accuracy discussed in Section 3.2.3, the orders of accuracy are defined using the L2 norm of the solution changes. Order of accuracy $p_{SN}^{(1)}$ of the first term in the error expansion is given by

$$p_{SN}^{(1)} = \frac{\ln(\left\|\mathsf{e}_{32_{SN}}\right\|_2 / \left\|\mathsf{e}_{21_{SN}}\right\|_2)}{\ln(r_{SN})} \tag{C.22}$$

which is used to provide a pointwise error estimate $\mathsf{d}_{RE_{SN_1}}^{*(1)}$ of the leading term in the error expansion

$$\mathsf{d}_{RE_{SN_1}}^{*(1)} = \frac{\mathsf{e}_{21_{SN}}}{(r_{SN}^{p_{SN}^{(1)}} - 1)} \tag{C.23}$$

Similarly, the equations for the first two terms in the error expansion are

$$p_{SN}^{(1)} = \frac{\ln[\mathrm{a}_{SN} - \mathrm{b}_{SN}]}{\ln(r_{SN})}$$

$$p_{SN}^{(2)} = \frac{\ln[\mathrm{a}_{SN} + \mathrm{b}_{SN}]}{\ln(r_{SN})} \qquad \text{(C.24)}$$

where

$$\mathrm{a}_{SN} = a_{SN} / b_{SN}$$

$$\mathrm{b}_{SN} = -\sqrt{c_{SN}} / b_{SN}$$

$$a_{SN} = \left\| e_{21_{SN}} \right\|_2 \left\| e_{54_{SN}} \right\|_2 - \left\| e_{32_{SN}} \right\|_2 \left\| e_{43_{SN}} \right\|_2$$

$$b_{SN} = 2\left(\left\| e_{21_{SN}} \right\|_2 \left\| e_{43_{SN}} \right\|_2 - \left\| e_{32_{SN}} \right\|_2^2 \right)$$

$$c_{SN} = -3\left\| e_{32_{SN}} \right\|_2^2 \left\| e_{43_{SN}} \right\|_2^2 + 4\left\| e_{21_{SN}} \right\|_2 \left\| e_{43_{SN}} \right\|_2^3 + 4\left\| e_{32_{SN}} \right\|_2^3 \left\| e_{54_{SN}} \right\|_2$$

$$- 6\left\| e_{21_{SN}} \right\|_2 \left\| e_{32_{SN}} \right\|_2 \left\| e_{43_{SN}} \right\|_2 \left\| e_{54_{SN}} \right\|_2 + \left\| e_{21_{SN}} \right\|_2^2 \left\| e_{54_{SN}} \right\|_2^2$$

$$\mathrm{d}_{RE_{SN_1}}^{*(2)} = \frac{r_{SN}^{p_{SN}^{(2)}} e_{21_{SN}} - e_{32_{SN}}}{(r_{SN}^{p_{SN}^{(2)}} - r_{SN}^{p_{SN}^{(1)}})(r_{SN}^{p_{SN}^{(1)}} - 1)} - \frac{r_{SN}^{p_{SN}^{(1)}} e_{21_{SN}} - e_{32_{SN}}}{(r_{SN}^{p_{SN}^{(2)}} - r_{SN}^{p_{SN}^{(1)}})(r_{SN}^{p_{SN}^{(2)}} - 1)} \qquad \text{(C.25)}$$

Recall from Appendix B that the conditions for applying equation (C.24) were that (i) $\mathrm{a}_{SN} \pm \mathrm{b}_{SN} > 1$ (related to monotonic convergence) and (ii) $c_{SN} \geq 0$ (related to the solutions being in the asymptotic range). It was found that condition (ii) was not satisfied for grid sizes $\mathrm{D}x > 3.124x10^{-4}$ for the first-order scheme.

Tables C.2 and C.3 show the orders of accuracy for the first-order and second-order methods, respectively. For the first-order method, the three-grid estimate approaches the theoretical rate $p_{th}^{(1)} = 1$ from below as the grid is refined, whereas the five-grid estimate approaches the theoretical rates $p_{th}^{(1)} = 1$ and $p_{th}^{(2)} = 2$ from above and below, respectively, as the grid is refined. For the second-order method, the solutions are in the asymptotic range even on the coarsest grids, although $p_k^{(2)}$ is larger than $p_{th}^{(2)} = 4$.

Figure C.2 compares the exact comparison error $E=A-S$ to the three-grid error estimate and the single step size error estimate for d_{SN}. For the first-order method, the three-grid estimate is relatively poor especially for the coarser grids, whereas for the second-order method the three-grid estimate is close to both E and d_{SN}. Figure C.3 compares E to both the three- and five-grid error estimates. For the first-order method, the three-grid estimate is less accurate than the five-grid estimate especially for the coarser grids, whereas for the second-order method both the three and five grid estimates are accurate. The results show that the higher-order terms are more important for lower-order methods on coarser grids.

C.4 Correction Factors, Uncertainties, and Verification

Methods are needed to account for the effects of higher-order terms for practical application of RE when solutions are outside the asymptotic range. Figure C.4a and C.5a compare the true error E to the three-grid error estimate $\mathsf{d}^{*(1)}_{RE_1}$ vs. step size at one spatial location ($x=1$ for the first-order method and $x=1.1$ for the second-order scheme since maximums of numerical error occur there) and for the first- and second-order methods, respectively. The three-grid estimate accurately estimates the true error E for smaller step sizes, but over predicts E for larger step sizes. Closer examination reveals that the reason equation (C.23) over estimates the error is due to the fact that equation (C.22) under estimates the order of accuracy, as also shown in the figures and previously in Table C.2. Therefore, one approach is to correct the three-grid estimate by a multiplication correction factor, which accounts for this deficiency, i.e.

$$\mathsf{d}^{*}_{k_1} = C_k \mathsf{d}^{*(1)}_{RE_1} \tag{C.26}$$

Two definitions for C_k were investigated. The first is based on equation (B.8) substituted for the left-hand side of equation (C.26) and solving for C_k, but replacing $p_k^{(1)}$ from equation (B.9) by an improved estimate $p_{k_{est}}$ based on the modified equation $p_{k_{th}}$ (or in the general case on solutions for simplified geometry with similar step size expansion)

$$C_k^{(1)} = \frac{r_k^{p_k} - 1}{r_k^{p_{k\,est}} - 1} \tag{C.27}$$

Similarly, the second is based on equation (B.15), but replacing $p_k^{(1)}$ and $p_k^{(2)}$ from equation (B.16) by improved estimates $p_{k_{est}}$ and $q_{k_{est}}$

$$C_k^{(2)} = \frac{(\mathsf{e}_{23_k}/\mathsf{e}_{12_k} - r_k^{q_{k\,est}})(r_k^{p_k} - 1)}{\left(r_k^{p_{k\,est}} - r_k^{q_{k\,est}}\right)(r_k^{p_{k\,est}} - 1)} + \frac{(\mathsf{e}_{23_k}/\mathsf{e}_{12_k} - r_k^{p_{k\,est}})(r_k^{p_k} - 1)}{(r_k^{p_{k\,est}} - r_k^{q_{k\,est}})(r_k^{q_{k\,est}} - 1)} \tag{C.28}$$

Figure C.4a and C.5a also compare E with error estimates based on equation (C.26) with equation (C.27) and (C.28). Both estimates are closer to E than the uncorrected three grid estimate $\mathsf{d}^{*(1)}_{RE_1}$, but for coarser grids $C_k^{(1)}$ is somewhat too small and $C_k^{(2)}$ is slightly too large. Figure C.4b and C.5b show the same trends, but directly compare the exact correction factor $E/\mathsf{d}^{*(1)}_{RE}$ to equation (C.27) and (C.28). In this case, $C_k<1$ indicates that the leading-order term over predicts (higher-order terms net negative) the error. However, for the general case, C_k is equally likely to be <1 or >1 depending whether the order of accuracy is approached from below or above, respectively. $C_k >1$ indicates that the leading-order term under predicts (higher-order terms net positive) the error. Thus, for the general case the correction to the leading-term error estimate is equally likely to be positive or negative and can be used to define the simulation numerical uncertainty.

Equation (C.26) is used to estimate U_k or d^{*}_k and U_{k_C} depending on how close the solutions are to the asymptotic range (i.e., how close C_k is to 1) and one's confidence in equation (C.26). There are many reasons for lack of confidence, especially for complex

three-dimensional flows. Often, pointwise results are not uniformly convergent over all grid points (i.e., locally oscillatory or even divergent).

For C_k sufficiently less than or greater than 1 and lacking confidence, U_k is estimated, but not d_k^* and U_{k_C}. Figure C.4d and C.5d show that equation (C.26) can be used to estimate the uncertainty by bounding the error by the sum of the absolute value of the corrected estimate from RE and the absolute value of the amount of the correction

$$U_k = \left| C_k \mathsf{d}_{RE_{k_1}}^* \right| + \left| (1-C_k) \mathsf{d}_{RE_{k_1}}^* \right| \tag{C.29}$$

For C_k sufficiently close to 1 and having confidence, d_k^* and U_{k_C} are estimated. Equation (C.26) is used to estimate the error d_k^*, which can then also be used in the calculation of S_C [in equation (10)]. Figure C.4c and C.5c show that uncertainty in the error estimate can be based on the amount of the correction

$$U_{k_C} = \left| (1-C_k) \mathsf{d}_{RE_{k_1}}^* \right| \tag{C.30}$$

Note that in the limit of the asymptotic range, $C_k =1$, $\mathsf{d}_k^* = \mathsf{d}_{k_1}^* = \mathsf{d}_{RE_{k_1}}$, and $U_{k_C} =0$.

Uncertainty estimates enable a quantitative measure of verification for analytical benchmarks. A simulation is verified if equation (C.8) is satisfied and corrected simulations are verified if (C.9) is satisfied. Figure C.4c,d and C.5c,d indicate that the present solutions are verified at the chosen spatial location and at the levels (U_k, U_{k_C})=(15%, 7.5%) and (1.2%, .2%) for the first- and second-order methods, respectively.

Grid	number points	Dx
1	38401	7.8125 x 10^{-5}
2	19201	1.5625 x 10^{-4}
3	9601	3.125 x 10^{-4}
4	4801	6.25 x 10^{-3}
5	2401	1.25 x 10^{-3}
6	1201	2.5 x 10^{-3}
7	601	5 x 10^{-3}
8	301	1 x 10^{-2}
9	151	2 x 10^{-2}
10	76	4 x 10^{-2}

Table C.1. Grids

Dx	Estimate	$p_k^{(1)}$	$p_k^{(2)}$
6.25$x$$10^{-4}$	3 grid	0.94	-
	5 grid	1.06	1.46
3.125$x$$10^{-4}$	3 grid	0.97	-
	5 grid	1.01	1.73

Table C.2. Orders of accuracy for first-order method

Dx	Estimate	$p_k^{(1)}$	$p_k^{(2)}$
2.5$x$$10^{-3}$	3 grid	2.00	-
	5 grid	2.00	4.54

Table C.3. Orders of accuracy for second-order method
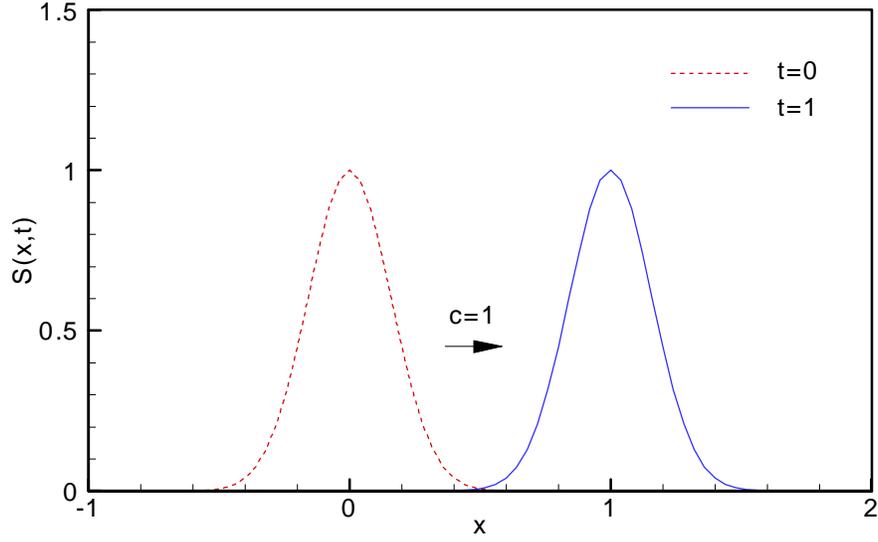
Figure C.1. Initial condition and exact solution for the 1D Wave Equation
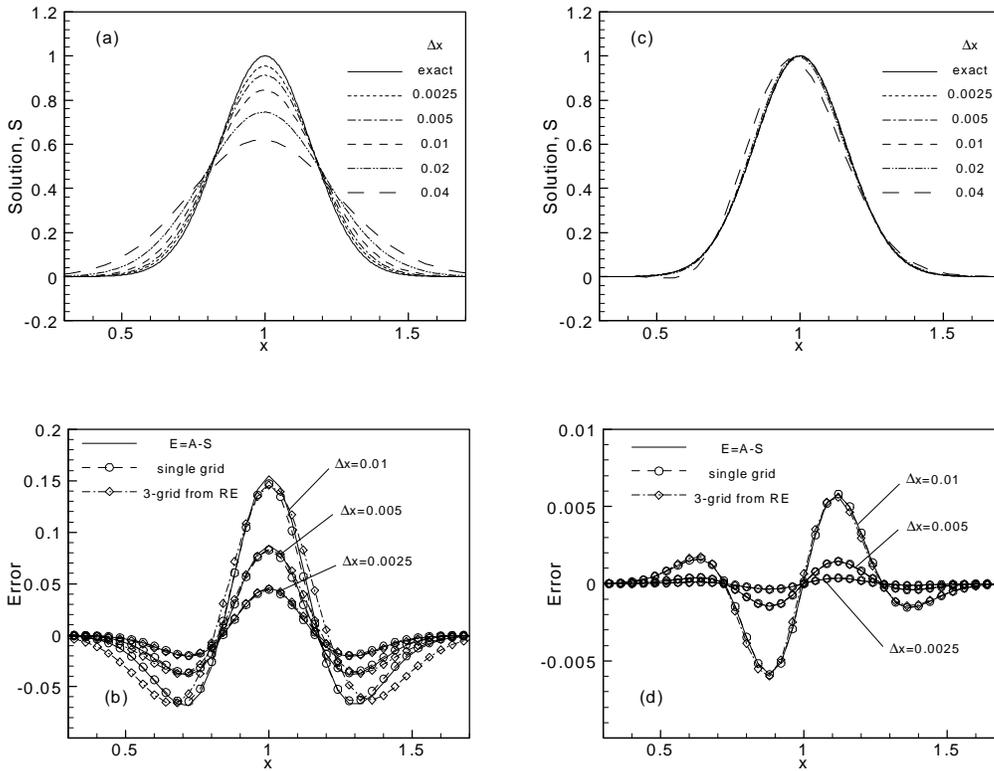
Figure C.2. Numerical solution S, true error E, single grid estimate d_{SN}^{*}, and three-grid estimate from RE for 1D wave equation at $t=1$: (a), (b) Euler explicit scheme (first-order) and (c), (d) second-order implicit scheme.
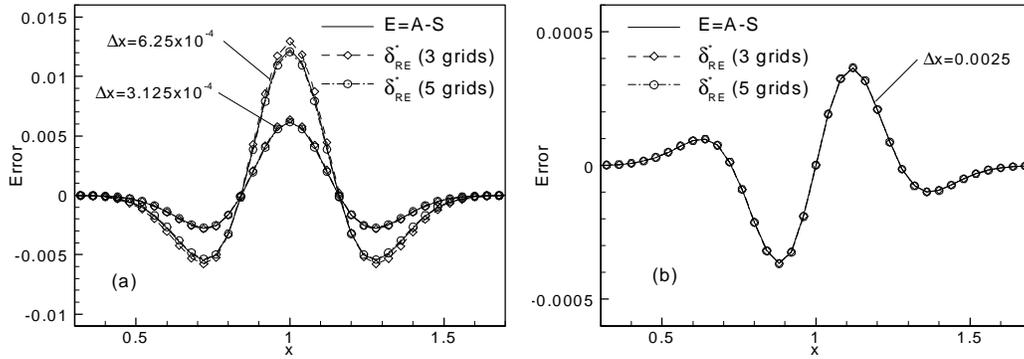
Figure C.3. Comparison of the true numerical error E to error estimates from RE using three and five grids for 1D wave equation at t=1: (a) Euler explicit scheme (first-order) and (b) second-order implicit scheme.
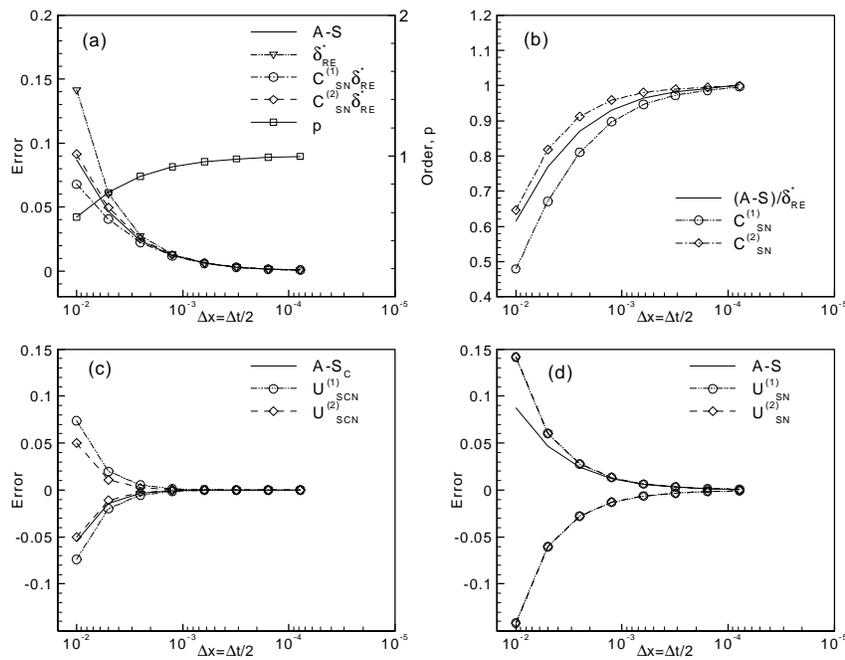
Figure C.4. Verification results for first-order numerical solution of 1D wave equation. (a) comparison of true error A-S to estimates from RE, (b) correction factor, and (c) comparison of A-S_C and $\pm U_{SCN}$, and (d) comparison of A-S and $\pm U_{SN}$.
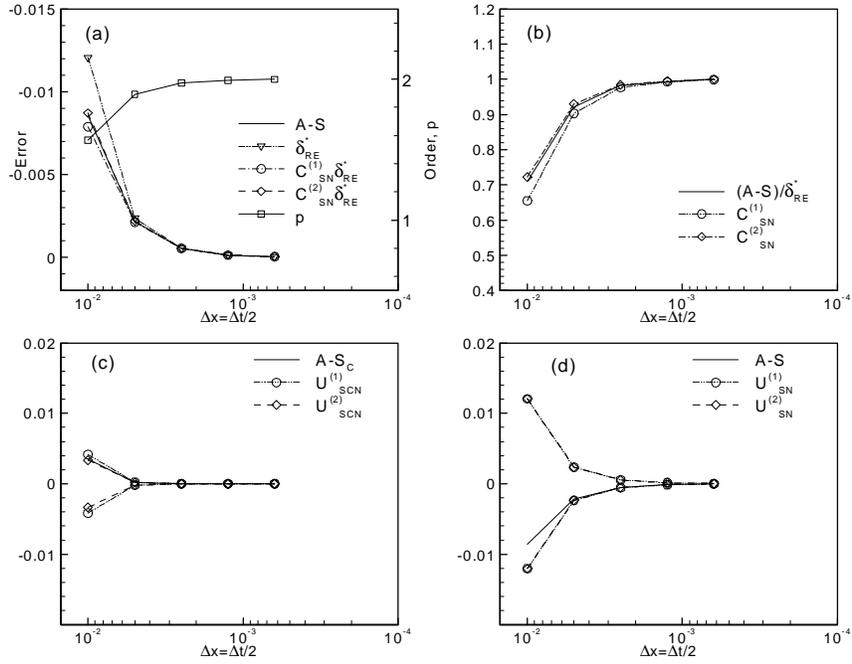
Figure C.5. Verification results for second-order numerical solution of 1D wave equation. (a) comparison of true error A-S to estimates from RE, (b) correction factor, and (c) comparison of A-S_C and $\pm U_{SCN}$, and (d) comparison of A-S and $\pm U_{SN}$.