AIAA JOURNAL
Vol. 36, No. 5, May 1998

Toward Establishing Credibility in Computational Fluid Dynamics Simulations

Arthur Rizzi*
KTH Royal Institute of Technology, 10044 Stockholm, Sweden
and
Jan Vos†
Swiss Federal Institute of Technology, 1015 Lausanne, Switzerland

Essential steps toward establishing credibility in computational fluid dynamics (CFD) simulations are outlined, and a vision for the process of systematic collaborative validation that is open to public scrutiny via the Internet is suggested. It begins with an exposition of the elements of CFD simulations and reviews protocols useful for establishing credibility. The various sources of uncertainty in CFD, which include the skills of the user, are presented. Lessons learned from collective verification and validation exercises done in the past are surveyed and lead to our suggestion for a systematic validation process that requires the creation and use of a detailed flow taxonomy for the given application field. The code validator uses the taxonomy and an electronic database to carry out the validation process. This database archives but also gives easy access to trustworthy data and allows full public discussion and scrutiny of the information, comparisons, and hypotheses so that judgments and conclusions about the validation may be accepted or rejected by the scientific community at large. The taxonomy also is the basis on which the code user transfers credibility from previously validated generic flow cases to the simulation at hand.

I. Introduction

COMPUTATIONAL fluid dynamics (CFD) is an emerging technology where it is now possible to apply full Reynolds averaged Navier–Stokes (RANS) CFD methods to geometrically realistic three-dimensional problems on available supercomputers and advanced workstations. However, it is clear that CFD is not a mature technology in the way, for example, that computational structural mechanics (CSM) is today.[1] In CSM, the availability of adaptive finite element methods that apply to general geometries and that require only modest computing resources to produce timely results has led to its pervasive use for structural design, including optimization. In CFD today, the situation is significantly different. Here, the need to deal with complex physics, multiple length scales and turbulent and/or reacting flows results in significantly greater computer resource requirements and has led to large numbers of specialized CFD codes limited to specific problems. This, coupled with the fact that these codes have most often not been validated for the class of problems being considered, has led to limitations that are not clearly defined. Consequently, the use of CFD for design work requires highly skilled CFD specialists to choose the physical models for a specific application, to execute it, and to extract the useful information. These difficulties seriously impede the acceptance of CFD in the product design and engineering community. Reducing the uncertainties of CFD simulations would ease this impedance and, therefore, is clearly of major importance.

Of course, there has been substantial effort in the recent past at assessing the capability of CFD codes for modeling a variety of flow problems.[2–6] However, much of the effort has generally been focused on issues of numerical accuracy and the ability of research and pilot codes to predict detailed flow physics for relatively simple problems and geometries. Although such studies were obviously necessary for developing this class of CFD codes, they, by themselves, were not sufficient to determine an accepted level of credibility. For example, rarely was the level of credibility of complex simulations established. Also, there is no general agreement on what is an effective standard for assessing the quality of CFD simulations. Although what needs to be done to enhance credibility is highly subjective and uncertain, it must include assessments of the accuracy of predictions of the performance data used by designers on, for example, classes of aircraft configurations they are developing. Performance data include such things as lift, drag, moment, maximum lift and loads, flap effectiveness, stall prediction, buffet onset, internal flow losses, etc. Although many studies along these lines, e.g., Ref. 5, have appeared in the literature, these tended to be ad hoc, one-of-a-kind-type investigations without broad implications. For example, the fundamental question of how well available RANS codes and turbulence models predict the maximum lift of a simple transport aircraft remains largely unanswered.

The present work is an attempt to outline essential steps toward establishing credibility in CFD and suggests a vision for the process of systematic collaborative validation that is open to public scrutiny via the Internet. In developing this vision, which is seen from a European perspective, we draw on concrete examples, mainly from the aerospace field, to illustrate concepts and principles. The paper begins with an exposition of the elements of CFD simulations and reviews existing protocols that have been found useful for establishing credibility. The various sources of uncertainty in CFD, which include the skills of the user, are presented. Collective verification, calibration, and validation exercises done in the past are surveyed and commented on. Then our suggestion for a systematic and open validation process that uses a hierarchical database founded on a detailed flow taxonomy for the given application field is presented. Finally, we indicate how the validation process can be carried out through the use of dynamic electronic databases that archive but also give easy and rapid access to trustworthy data and allow via the Internet full public discussion and scrutiny of the information, comparisons, and hypotheses so that judgments and conclusions may be accepted or rejected by the scientific community at large.

II. Elements of CFD Simulations

CFD simulations are used in research to aid in understanding the fluid dynamics, in technology development for developing theoretical and simulation models, in engineering for supporting design processes and wind-tunnel test or flight-test activities, and in making policy decisions. At the heart of all of these activities and

Presented as Paper 96-2029 at the AIAA 27th Computational Fluid Dynamics Conference, New Orleans, LA, June 17–20, 1996; received Oct. 22, 1996; revision received May 28, 1997; accepted for publication May 30, 1997.

*Professor, Department of Aeronautics.
†Senior Scientist, Fluid Mechanics Laboratory.

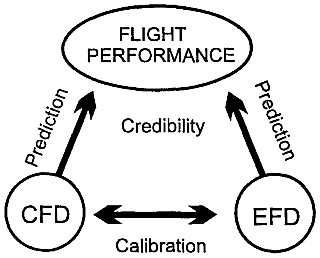


Fig. 1 Prediction of performance and calibration by CFD and EFD.

really the end goal of CFD simulations is to make predictions and gain knowledge.

A. Performance Prediction by Simulation

In the aerospace field engineers need methods to predict the flight performance of the air and space vehicles that they are designing. Their two main tools for this prediction (Fig. 1) are CFD simulations and ground-based testing and measurements [experimental fluid dynamics (EFD)]. Not all performance quantities can be predicted by both tools, and when they can be there are usually differences in the cost and accuracy of the predictions. But if we leave cost aside, when both CFD and EFD deliver a performance prediction, then the two results can be compared. This forms a first basis for assessing the quality of the prediction, and comparative agreement gives the engineer confidence in using the result. The engineer can then go on to the next level and calibrate the simulation model to improve the predictions, but this, of course, limits the generality of the model. Some people claim that this last step is overly restrictive, but we hold that, in the absence of a universal turbulence model, calibration remains a practical necessity and, therefore, is a part of our simulation paradigm.

The performance quantities are forces, moments, heat loads, etc. The credibility of performance quantities depends on the credibility of the simulation, which in turn depends on the credibility of the modeling of phenomena such as turbulence, relaminarization, retransition, catalyticity, chemical kinetics, combustion, shock interactions, shear-layer mixing, separation, secondary flows, etc. Whether the simulation is ultimately satisfactory is a function of the accuracy of the prediction and its cost in terms of work hours, calendar time, user skills, computer resources, and computational and modeling complexity.

B. Simulation Models

Following Mehta,[6] the simulation model is the physical (conceptual) model manipulated with computational models and a set of computational or logical associations (numerical solution procedures) for creating simulations based on these conceptual and computational models. Both computational models and fluid-dynamics models essentially have features (structures) and parameters.

The conceptual model consists of global and local fluid-dynamics models together with initial and boundary conditions. Examples of global and local models are the Navier–Stokes (NS) equations and models for turbulence and transition from laminar to turbulent flow, respectively. Examples of computational models are structured or unstructured finite volume methods, Runge–Kutta time-marching schemes, etc. The order of the discretization scheme also sets an important feature to the simulation model. The number of grid points and the convergence criterion of an iterative scheme and any data related to computational processes are examples of computational model parameters.

III. Protocols for Establishing Credibility

For reasons of completeness but primarily to establish the context for the main thrust of the paper, we review here some of the concepts and terminology useful for the discussion of the credibility of CFD simulations. There is a good deal of variation in the terminology used in the literature for describing the different stages in ascertaining the fitness for use of a code. It is not our purpose to add to this body of terminology nor to further refine or expound upon it. Much of our usage in what follows is rooted in the terminology and the guidelines suggested by Mehta.[6] Any one of the terms *verification*, *calibration*, *validation*, or *certification* can be found in the literature relating to CFD in discussions of the accuracy or credibility of CFD simulations. However, these terms, which denote processes, and the words from which they derive mean different things to different people, e.g., see Refs. 4 and 7–9. For example, to Mehta[6] calibration is only the process of obtaining correction factors, whereas to us this process is a part of our simulation paradigm, as noted earlier.

A. Definitions

The term *verification* addresses the accuracy of the computational model, the terms *calibration* and *validation* are concerned with the suitability of the physical model, and the term *certification* encompasses both the computational and the physical models. The following definitions and usage are intended only to facilitate our discussions, mostly concerning the verification and validation processes.

Verification is the process that demonstrates the ability of the computer program to solve the specific set of governing equations and boundary conditions posed to the computer by the simulation model and its input. It establishes the level of accuracy of the solution and the sensitivity of the results to parameters appearing in the numerical formulation through purely numerical experiments involving grid refinement studies and comparison to available exact solutions. It does this for all output variables and identifies the grid required to achieve a specified level of accuracy.

It is the role of the validation process to establish the degree to which the physical model is an accurate representation of reality from the perspective of its intended uses and the range of parameter space over which the model applies and thus build up some reliance and certainty in the predictive ability of the code. Validation is carried out by extensively comparing CFD simulations with trustworthy detailed experimental measurements and is usually done at three different levels for the different stages in the development cycle of a CFD code. In its earliest development stage, a research code is validated by comparison with benchmark experimental testcases of relatively simple flows exhibiting one single dominant flow feature over a simple geometry, e.g., a boundary layer over a wall or an airfoil. At the next stage, a pilot code is validated with component test cases involving flows containing two fluid phenomena that interact with each other over a component of the aerodynamic system, e.g., shock wave/boundary-layer interaction on a wing. At the final stage, a mature code should become a production tool and go into the hands of the engineering staff. A production code, therefore, is validated with system test cases that involve complex flows with interacting combinations of all features found in the full configuration, e.g., a complete transport aircraft for which the task at hand is to reliably predict design/performance data for configurations of interest over a specific operating envelope. Thus, the validation process at the highest stage relies on comparisons with experiments based on complete aircraft configurations and is focused on global design/performance data.

Calibration is the process of tuning or calibrating a code with a particular fluid-dynamics model to improve its prediction of global quantities for realistic geometries of design interest. Calibration, however, needs to be carried out for several test cases, each with different physical features, e.g., separation, no separation, low Mach numbers, transonic Mach numbers, high Mach numbers, vortex flows, flows with high streamline curvature, etc. In fact, this comes down to defining a class of restricted flow phenomena, and this is an important step toward validation. A good example of calibration is the Degani–Schiff modification of the Baldwin–Lomax turbulence model, where in the direction normal to the wall the first local maximum of the vorticity is used instead of the overall global maximum as called for by the theory. This tuning or constraining the model is appropriate for cases of leading-edge vortices over delta wings, but the generality of the implementation of the model is lost. The result is that the simulation is validated for only a limited class of flow problems because a confirmation of reality with a calibrated model may be fortuitous when it is used for simulations under conditions different from those holding during the calibration. Calibration is not a process determining the level of accuracy or credibility but a process of setting parameters or obtaining correction factors.

Certification is the process having perhaps the widest diversity in definition. As defined herein (following Ref. 6), it is the process of evaluating software in terms of its logic, conceptual and computational models, procedures, and rules and documentation and in terms of the simulations derived from it to ensure compliance with specific requirements for the intended uses of the software. In part, it is concerned with programming matters (logic checks and programming style), documentation, and quality assurance issues, e.g., rerunning the suite of certification test cases after each new version release to be certain that no new errors have been introduced into the previously certified version. But to a great extent certification also concerns the credibility of the simulations. After the successful completion of verifying the numerical model and validating the physical model, most of the requirements for certification of the software are addressed. The remaining part of these requirements deals with the rules for use of the simulation model and its documentation. For example, the same code or simulation model may or may not result in a credible simulation, depending on how it is used, raising the issue of end use and user skills (more on this subsequently). A satisfactory certification guarantees that the software complies with its specified requirements and is acceptable for operational use. In the words of Mehta,[6] certification assesses whether the right things are done and whether they are done right.

As an example of this at the coding level, consider a typical programming implementation. In the Baldwin–Lomax turbulence model one has to specify where the inner layer ends and the outer layer begins. One way that perhaps is more physically correct is to specify a flipping function, e.g., where $y+ = 35$, to determine the boundary. We have done this in our NS code, NSMB,[10] but have found that this inhibits convergence on the medium mesh for a transonic wing case because it causes a limit cycle in the determination of the turbulence quantities. Another choice of layer-transition function is an exponential, and this works much better because it gives a smooth transition. Both are deemed correct, but the latter is clearly more desirable and better fulfills the requirements of the software. Sorting out this type of question is part of the work towards certifying CFD software. The paper by Melnik et al.[1] presents a comprehensive overview of the certification process in its entirety, illustrated by concrete examples.

B. Teamwork

As an engineering activity, CFD simulations involve the interaction of three distinct teams. Software engineers, mathematicians, and physical model developers make up the first team. Experts with a strong background in the physical models, who carry out the calibration and validation work, are the second team. Users who are experts in neither the code or modeling but are application specialists make up the third team.

Of the four steps to build credibility, verification is seen as the responsibility of the first team, the code developers. The work of the second team and the work of the third team are interrelated because calibration and validation are specific to a particular design application and user community. It is the task of the second team to calibrate a code and through validation to determine its range of applicability, which then has to be transferred to the end users through the certification process. Validation and calibration are related. In calibration one improves the results of a validation at the expense of a general validation because one modified or specialized the physical model to make it work for a certain type of generic flow. In validation one identifies all of the generic flows types that make up the complex flow and then chooses the appropriate model to make the predictions. If the resulting simulations then meet the requirements of the end user, the software can be certified and adopted for operational use. Thus, certification is the work of all three teams. The reasoning behind this approach is inductive, and therefore its credibility hinges on the validation and certification process. The key activity in this work, which is also the most demanding and expensive, is validation/certification. A strong case exists for collaborative validation, involving a number of different codes and at least two rounds of analysis. It is likely that a range of studies would be needed to satisfy both model developers and industrialists.

IV. Sources of Uncertainty in CFD

All modeling involves approximations of one kind or another. These approximations definitely set the accuracy of the predictions. The outcome of a validation is meaningless without any indication of the effects of the numerical errors on the physical model. Specifically, the sensitivity of simulations to discretization errors must be established, in the simplest case by comparing two simulations with different grid resolutions. Ideally, unambiguous validation can be achieved only if the computational errors are separated from the errors in the physical modeling. Issues related to the interaction of these two sources of errors have been of constant concern to the CFD and turbulence modeling communities because fortuitous cancellation of errors can lead to a very erroneous conclusion about validation. A good example is the simulation of a shock wave interacting with a boundary layer using, e.g., the Baldwin–Lomax turbulence model. At one level of grid resolution the simulation compares rather well with the experiments in terms of shock wave position, and one might conclude that it is validated. However, a more refined-grid simulation turns out not to compare as well; the shock position has changed. The reason for this is that the turbulence model does not have a finite relaxation time and is not very accurate for this case. The comparison is better in the coarser grid simulation because the discretization errors fortuitously canceled out the errors in the turbulence model. Instead of improving the comparison, refining the grid reduces the discretization error, and the cancellation becomes less and, thus, degrades the comparison.

Examples of uncertainty arising from lack of numerical accuracy are 1) discretization, which includes poor shape definition of the configuration due to design changes, grid quality, truncation errors, and artificial viscosity; and 2) errors introduced by postprocessing of simulated results. It would be very desirable for the code developers to build error indicators into their programs. These could range from very simple systems, such as checks on conservation of mass flow, to extremely sophisticated systems, invisible to the user, setting firm bounds on the final solution.

Mehta[6] lists various causes of physical modeling uncertainty: 1) The phenomenon under investigation is not thoroughly understood; 2) parameters used in the model are known but with some degree of uncertainty; 3) appropriate models are simplified, thus introducing uncertainty; and 4) an experimental confirmation of the model is not possible or is incomplete.

A. End Use and End-User Skills

Finally, the level of validation and the degree of certainty depends very strongly on its end use and the end user. The more thorough the validation is, the less uncertain the results will be. But thorough validation studies require much time and long computer runs, hence leading to high costs. This is particularly critical in industry, and thus tradeoffs have to be made. It is the end use that tempers the choice of simulation model, and in industry this usually is in the direction of just enough certainty for the problem at hand. The following example illustrates this point. In many industrial applications the resolution of the boundary layer is not always accurate due to restrictions on the number of gridpoints that can be used and due to restrictions on the mesh generator system. Typically, the engineers at Aerospatiale make a three-dimensional NS simulation over a wing–body–nacelle configuration using more than 1×10^6 grid points, but it is estimated that to resolve all viscous effects in the flow at least 10 times more grid points are needed. However, the 1×10^6 grid point simulation takes at least 30 CPU h on a current vector computer, and the elapse time is on the order of several days. Therefore, the compromise is made in the direction of expediency.

The end user of the code needs to be aware of these compromises and needs to know whether a code can predict the design parameters of interest, whether it can predict trends (derivatives) of these parameters, or whether it can only describe the flowfield characteristics in general. For complex simulations, the end user should decide on 1) minimum physical modeling needed to predict the quantity of interest (Euler flows, RANS, turbulence model, etc.); 2) numerical issues, for example, numerical scheme, numerical dissipation, Courant–Fredrichs–Lewy numbers, and boundary conditions; and 3) grid clustering to get the quantities of interest to the required level of accuracy.

When the end user lacks these skills, then the accuracy obtained in (complex) CFD solutions often falls short of the accuracy expected based on simpler benchmark validations.

CFD calculations are made to simulate reality by the introduction of a numerical model. Because the direct numerical simulation of turbulence (DNS) for complex high-Reynolds-number flows is not feasible and will not be possible in the coming decades due to the large computer resources required, turbulence is approximated by the introduction of a turbulence model. The more the end user knows about the numerical and physical models, how to use them, and how to select the grid and numerical parameters, the more credible the result will be.

The complex nature of both the physical understanding and the mathematical expertise required raises questions about the training schemes needed for the teams of developers, validators, and users. One suggestion is to consolidate the expertise in modeling and numerics, which currently resides with researchers and code developers, together with the knowledge built up in calibration and validation, into a book, or even better a multimedia training aid, of best CFD practice for end users.

B. Estimating Credibility

The credibility of a simulation model is established as follows. First, the computational uncertainties are determined with verification using sensitivity analysis. If computational uncertainties are negligible, then the uncertainties resulting from the conceptual model are obtained with validation. If possible, simulations are compared with measurements that incorporate known measurement uncertainties and for the same flow conditions, e.g., with wind-tunnel wall boundary conditions. If there is an acceptable level of agreement between the simulated results and the measurements, then these results, and hence the simulation model, are credible. If such measurements are not available, i.e., it is a prediction, then those simulation models that have been demonstrated to be credible, e.g., through earlier calibration and validation processes, should be shown to be applicable to the flows of interest by reference to a flow taxonomy (see Sec. V). Whether the uncertainties associated with these calibrated models are transferable to the problem of interest, of course, needs to be investigated. If these conditions are met in some sense, then simulations of these credible models are compared with those from the simulation model of interest and uncertainties are determined.

V. Uncertainties in EFD

Although the primary focus of this paper is the credibility of CFD simulations, uncertainties are also inherent in EFD, and as mentioned in the preceding section, these uncertainties can undermine the credibility of the physical modeling used in CFD simulations. Mehta[6] points out that uncertainties arise in both the fluid dynamics itself and in the measurement taking. The fluid-dynamics uncertainties come about when testing is done under conditions other than the operating conditions of the fluid-dynamics system. For example, the ground-based facilities may manifest phenomena other than or in addition to those likely to occur in flight. In the process of taking measurements, either ground based or in flight, there are interference uncertainties and data uncertainties. An example that we encountered illustrates this point. We recently carried out a routine validation check on our Reynolds-averaged NS code by comparing the computed results with the measured results reported for the MBB-AVA pilot model in Ref. 11. The comparison showed that the suction peak computed on the upper wing surface was about 10% lower than the observed value. Additional computations and careful checking of our results did not lead to an explanation for the discrepancy. We then contacted one of the experimentalists, who told us that during the experiment the model underwent a substantial bending and permanent deformation of the right wing due to the lift load. This information was not reported in Ref. 11, and so the data are incomplete. Even if nothing else, the deformation would change the effective angle of attack and introduce some uncertainty in the data and make its use for comparison ambiguous.

Standards are now available and should be adhered to for quantifying measurement errors and assessing wind-tunnel data uncertainties,[12–14] and Oberkampf and Aeschliman[15] describe a method of statistically estimating total experimental measurement uncertainty in wind-tunnel testing. For a more detailed discussion on the uncertainties in EFD, we refer the reader to Ref. 16.

VI. Previous Collaborative Verification/Validation Exercises

Results computed with a large computer code cannot be judged in prepublication review to be correct in the same way that a paper presenting an analytical result can be judged. The typical reviewer can devote at most a day to a given review, and this is grossly insufficient to unravel a large program, even if a listing of the source code is provided. About all the reviewer can do is examine the general statements provided about the code and see whether agreement with the data is obtained. Therefore, the usual criteria for computational results have been of necessity comparison with one or more sets of either previously accepted simulations or experimental data. Out of this situation arose the idea of a workshop, a collective learning process where a group of researchers actively working on a sharply defined topic meet to discuss in detail their problems and experiences and to make direct comparison and critical evaluation of their data. We review a few such collaborative activities in the past for the purpose of pointing out how their results have been archived in static databases and of contrasting this with the possibilities today of archiving with dynamic databases on the Internet.

A. Collaborative Workshops

During the 1970s and 1980s the Gesellschaft für Angewandte Matematik und Mechanik (GAMM) Specialist Group for Numerical Methods in Fluid Mechanics organized a number of workshops that illustrate aspects of the verification, calibration, validation, and certification processes as carried out in collaborative exercises.[17–20] When reading about them, keep in mind the evolving nature and latitude in usage of the terms, as pointed out earlier. One of these workshops dealt with the problem of inviscid steady transonic flow, and its goal was to assess the numerical accuracy of numerical solutions to the inviscid equations.[17] It was perceived that the best way to determine this was by collective comparison of the results obtained from the computational methods that were in use at the time. A number of well-chosen test problems were defined, and solutions with mesh refinement were obtained. The purpose was to gain some insight into the methods and ultimately to yield a set of definitive solutions of inviscid transonic flows, which would stand as benchmark cases. In some sense the goal was reached, but with hindsight perhaps the most lasting result of this workshop was the first report by Jameson (Ref. 17, Appendix C) that the solution to the transonic full potential equation in conservation form is nonunique. Jameson followed up on this with thorough tests, which were then repeated and confirmed by other workers. That the solution to transonic potential flow is nonunique is now generally accepted by all because scientific truth is established through such a process of public scrutiny and repeatability of results. Collaborative workshops foster such activity.

A significant validation/certification exercise was the Viscous Transonic Airfoil (VTA) Workshop[21] sponsored by the AIAA Fluid Dynamics Technical Committee for the purpose of validating viscous transonic airfoil computations over a range of geometries and flow conditions. The primary objective was to gauge the predictive capability of the methods, and an additional purpose was to establish a database that could be used for future computer-code validation. But the database was not archived in machine-readable form; consequently, access was limited, the data became very static, and it is no longer in use. This is a very common outcome of databases from the past.

B. Stanford Paradigm for Validation

The early GAMM workshops dealt mostly with the verification process and functioned on a limited scale. The Stanford Conference of 1980–1981 was a landmark event that approached the problem of validation in a very comprehensive way and in the process established the paradigm for the collaborative validation process.[22]

Its three related goals were 1) to reach consensus in the research community on trustworthy data sets that can be used as input for modeling of turbulence in complex flows and as the basis for comparisons with computations, 2) to create a machine-readable data library to hold the data selected as trustworthy in standard normalized form, and 3) to compare current computational results for a set of basic test cases of turbulent flows.

The conference was a very professional attempt to give an overview of the state of the art in computing complex turbulent flows in 1981 using the experimental database established for this purpose in 1980. The paradigm still stands today as complete in the sense of providing all elements necessary for understanding the state of the science and for carrying out CFD validation. These elements are taxonomies that organize the flows, methods of modeling, and numerics; clarification of experimental data and computational data via thorough review processes; all computational results compared with experimental data and with other computed results, case by case; examination, analysis, and discussion of the comparisons; an overall synthesis of the findings by distinguished experts; and storage of all data in specified machine-readable format.

Perhaps the element that distinguishes the Stanford Conference most is recognition of the importance of archiving all of the data for future use. Without a system such as the data library, the personal knowledge and files of the individual researchers would have been lost to accurate recapture over time. Public access to the data is a very essential issue because information and hypotheses become science only when they are subjected to full public discussion and scrutiny and become accepted by the large majority of workers in the relevant field.

VII. Hierarchical Databases for Systematic Validation

In all of the exercises discussed and in those pointed to in the references, when a database does exist, it is static in the sense that it is not updated and cannot be interrogated using graphical tools. If no machine-readable database exists, then the data are found only in written reports, and if you use them today, there is additional uncertainty in extracting these data from graphs. The complete data may not even be reported, and so there is a loss. The test cases typically have been for simple configurations (two-dimensional airfoils, wedges, cones) and three-dimensional wings and bodies. Even for generic flow cases for calibration, these are too simple. In some cases the input description is incomplete and the output format for the results have been specified imprecisely. Clearly, there is a need for more advanced databases that are accessible and updatable electronically and that offer interrogation with modern graphical tools. Above all, we believe that such an electronic database must be hierarchical.

Consider the situation of a designer making a simulation to predict design quantities of a configuration for which no measurements are available. To what extent can the predictions be trusted, and what is the best way to build up the credence? One way is to look at similar simulations for which the code (using the same models) has been validated by comparison with measurements. A taxonomy of the various flow classes then can help to select similar simulations, i.e., it places the large number of flow phenomena into some order and structure. This order can also serve as the foundation for an archival database because it makes the data more easily accessible. As already appreciated at the Stanford Conference, a taxonomy of flows is needed to identify and prioritize the needs for validation of CFD simulations and for analysis of experimental data of industrially relevant flows occurring in aerospace. Here we sketch what such a system for the categorization of flows encountered in the aerospace sector might look like and how it might be used. Let us emphasize that this is nothing more than a sketch for the purpose of illustrating the concept. We select just a few flow categories that come to mind from external airframe problems, with obvious omissions from propulsion, rotorcraft, and many other areas. Ours is not a comprehensive breakdown of all flow categories, which would go beyond the scope of this paper.

A. Categorization System

The widespread use of CFD simulations and the need for their calibration and validation, together with the diverse range of application problems to which they are put, demand a hierarchical approach to the creation of a taxonomy of flows appropriate to the needs of the aerospace sector. The following might serve as appropriate levels of the hierarchy: A, flow regime—that part of the flight envelope that is under study; B, system—the full aircraft configuration or system that defines the application; C, component—an individual component or subsystem experiencing the flow; and D, flow feature—the elements of physical behavior of the fluid.

Of course, the creation of a comprehensive taxonomy of flows within the aerospace sector is a huge task and is not attempted here. The following is intended only to describe a systematic framework for the purpose and to sketch what the classification might look like, mainly to indicate how the system might be used.

Flow Regimes

At different points along its flight envelope, an aircraft can experience flow regimes ranging from low-speed flow, steady flow, unsteady flow, transonic flow, buffet conditions, high lift and stall conditions, to supersonic or hypersonic flow.

Aircraft Configuration and Components

These, the second and third levels of the hierarchy, are defined by straightforward but large sets and subsets of categories that identify the range of scope of aircraft types (civil, transport, military) together with integral subsystems and components that exploit or experience the flow of fluid.

The following example illustrates the distinctions between hierarchy levels A, B, and C in consideration of a commercial transport plane in straight-and-level flight. The flow regime in which the aircraft finds itself is steady transonic flow. The configuration (level B) is a wide-body aircraft, whereas the component (level C) is a large-aspect-ratio wing. Of course, the flow over the main wing is only one of a number of flows associated with the configuration, of which each is characterized by different physical phenomena and may be subjected to separate analysis or investigation and should therefore be recognized separately at level C.

Flow Feature

This, the final level in the classification hierarchy, is the stage at which the scientific aspects of the flow are appreciated. The important aspects of the flow in question are identified in terms of the dominant features of the flow phenomena within it and the corresponding physics (and, perhaps, chemistry). Continuing with the example given, of the flow over the main wing, the flow can be specified as the three-dimensional, viscous, turbulent external flow of compressible air at flight Reynolds number, with moderate shock waves, complex geometry (including winglets), regions of flow separation, and significant wall roughness. These features, once recognized, can be used to prescribe the features of modeling needed to study the flow either computationally or experimentally. They also define the aspects of behavior for which a particular investigative or predictive tool must be validated for it to be applied reliably to that flow. Furthermore they point out the background and expert skills that a user needs to carry out the work, e.g., knowledge of separated flow and turbulence modeling.

B. Taxonomy of Flow Features

Table 1 represents a glimpse of what would be a preliminary attempt to categorize features of compressible flow in the aerospace sector, which are either fully turbulent or transition is known.

The definition of classes is never completely precise and is of necessity imperfect because single classes of features never occur isolated in a complex flow, but the categorization is an aid in establishing credibility and should be seen as such. The process is more difficult and less reliable without it.

VIII. Evolution to Dynamic Validation

The inductive process implies credibility to a predictive simulation through careful analysis and transfer of similar validated and credible simulations in a taxonomy of similar flow cases. For it to work properly, systematic databases of the generic flows in the taxonomy must be constructed, and the databases must be accessible. Furthermore the data must be dynamic, i.e., they must evolve. As new and better data comes along, it must be possible to correct and

Table 1 Example: compressible turbulent flow classes

I. Wall-bounded flow
I.1 Attached
I.1.a with pressure gradient
I.1.b without pressure gradient
I.1.c with shocks
I.1.d with or without suction
I.1.e with or without blowing
I.2 Separated
I.2.a bubble type
I.2.b with reattachment
I.2.c without reattachment
II. Free-shear-layer flow
II.1 Vortex flow
II.1.a delta wing leading edge
II.1.b forebody
III. Mixing layers
IV. Flows with confluence of two shear layers
IV.1 High-lift devices
IV.1.a off-body recirculation zones
V. Flow in a cavity
VI. Wake flows

refine the database after a suitable evaluation and acceptance of the proposed new entry by the database manager.

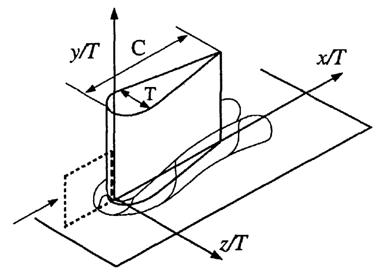
This concept is not new; it was spelled out at the Stanford Conference. All that has happened since then is that the technology for handling and storing the data has advanced, and we are now progressing toward the electronic meeting place where validation centers around a remote but electronically accessible database that is continually updated and refined. Just what constitutes an electronic database is not readily amenable to a concise definition. Loosely speaking, it is any collection of data meaningful for comparison that is on line and machine readable. Pointing to some examples gives a more precise idea. Consider the Stanford Conference. The data resided in a data library, which consisted of files (in ASCII format presumably) on a magnetic tape, a completely passive medium. In the European Hypersonic Database, which came out of the Antibes Workshops,[23] 1990–1993, the data, which included complete simulated flowfields, resided on a file transfer protocol (FTP) server. The data were stored in a specified format, and a visualization system was included to interrogate the results. The idea was good, but there were limitations in the funding to support the system. Next we discuss more recent databases using the World Wide Web (WWW) as the medium and with more modern graphical systems for data interrogation.

A. Validation with the SCIENCE Database

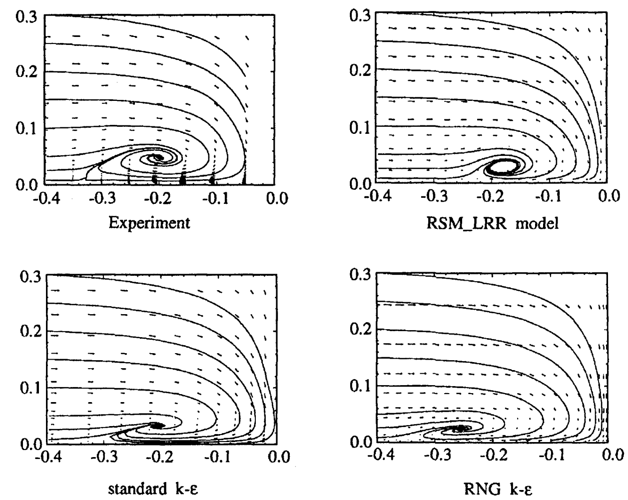
The research project Data Validation and Comparison in Fluid Mechanics, financed by the European Union (EU) Programme SCIENCE (for a progress report on the project, see Ref. 24), has as its aims to collect experimental and numerical data on turbulent flows, to check the data for their reliability and suitability for test cases, to set up test cases and perform calculations with various turbulence models, and to create a data bank from which the data can be accessed. The project work has yielded data for over 70 flows, and 15 well-documented test cases have been set up. Some of these test cases were the subject of a workshop where computors were invited to present and compare their simulations with the data in the SCIENCE databank at the University of Surrey accessed through http://fluindigo.mech.surrey.ac.uk/.

The prescription of all geometrical parameters, boundary conditions, the detailed experimental and, in one case DNS, data, and the instructions for presenting results were distributed to more than 110 groups all over the world, using the Internet facilities as much as possible: Almost all of the data files were sent via electronic mail, and the calculation results were submitted via FTP to a workstation set up at the site of the workshop. More than 500 MB of disk space were necessary to store the contributed result files, descriptions of numerical methods, and grid representations.

We summarize here the test case on wing–body junction with separation and present a typical result.[25] This case concerns the flow around a cylindrical airfoil mounted on a flat ground plate (Fig. 2a). The experiments are due to Fleming et al.,[26] who measured the



a) Flow configuration and plane where results are shown



b) Secondary flow velocity vectors and streamlines in symmetry plane in front of the wing

Fig. 2 Comparison of results for wing–body junction with separation.

streamwise and secondary flow velocities as well as the Reynolds stresses in several vertical planes. The boundary layer on the flat plate separates when it approaches the leading edge of the wing, and a horseshoe vortex, which sweeps around the wing in the junction corner, is formed (Fig. 2a). For this three-dimensional case, seven computers submitted results; six of them employed versions of the κ–ϵ model and one the Launder et al. Reynolds stress equation (RSE) model.[27] Five of the groups used wall functions, and two used the two-layer approach. In Fig. 2b velocity vectors and streamlines constructed from the simulations are given in the symmetry plane in front of the wing for three calculations, all using wall functions, together with the experimental results. It can be seen that the general flow picture is well captured by the various models but that the separation vortex is predicted too thin and, particularly with the RSE model, also too short. The experiment shows fairly high turbulent kinetic energy k in the vortex region with low values in the immediate stagnation region, whereas most models have the maximum of k more toward the corner and produce excessive k in the stagnation region in front of the leading edge. This is a well-known problem with the standard κ–ϵ model, but in this case the RSE model also suffers from it, albeit to a lesser extent. The best k distribution is obtained with the renormalization group (RNG) version of the κ–ϵ model. When sweeping around the wing, the strength of the horseshoe vortex is underpredicted by all models, the vortex being more diffuse and the higher k levels being too close to the wing. The models that do not yield excessive k in the stagnation region (RNG version and to a lesser degree RSE) are best. In the plane immediately downstream of the trailing edge of the wing, size and strength of the vortex are also underpredicted. Now k has two maxima, one toward the outer edge of the vortex and the other in the wake behind the trailing edge; this feature is picked up by all models and best reproduced by the RSE model, whereas the RNG model now gives too low k.

One observation of the workshop as a whole is that the Reynolds-stress and algebraic-stress models reproduce better details of turbulence quantities, but for the test cases considered these models are not consistently better than simpler two-equation models as far as the mean quantities are concerned.

B. Validation with HAEDB

The Hypersonic Aerothermodynamic Engineering Data Base (HAEDB) is a database system in which data from CFD simulations, data from experiments, flight data and synthesis documents, and data concerning re-entry vehicles are stored. Its purpose is to make data generated at different European organizations in the frame of European Space Agency (ESA) programs available in a complete and unified manner. Besides the archival function of the HAEDB, the results saved in the database can be used for validation of computational results with experimental data and/or flight data. Tools are provided with the HAEDB to find data saved in the HAEDB and to manipulate the data to generate, for example, synthesis documents. Access to the HAEDB is provided via the WWW; however, most of the data is not accessible from the public domain. In addition, the most sensitive data are saved off line and can be accessed only after intervention of the HAEDB manager.

The HAEDB database system goes beyond the database developed at INRIA during the series of Antibes workshops[23] because with the HAEDB one can search the data, compare results for different conditions, use data manipulation tools to generate automatically synthesis documents, and retrieve reports, pictures, and comments of experts that accompany the data.

The HAEDB is different from the European Research Committee on Flows, Turbulence, and Combustion Association (ERCOFTAC) Science database in the sense that not only are experimental data available but also computational data, along with the tools to manipulate the data. The ERCOFTAC Science database is more oriented toward a depository of data, where information is available and can be retrieved by a user. In addition to that function, the HAEDB incorporates expert knowledge, which increases the users' understanding.

Several levels of synthesis documents are used in the HAEDB. For example, the raw signal of pressure sensors as a function of time are saved in the HAEDB (off line in general), together with the value of the C_p. This C_p value is provided by the expert responsible for the experiment and can be used for comparison with computational results. Comparisons of computations with experiments yield a second type of synthesis documents.

The HAEDB system was used for the first time to generate the synthesis documents for the HALIS S4 test case of the ESA/ESTEC Workshop in March 1996, one of a series of workshops organized by the ESA.[28] As an example of the results produced with the HAEDB, Fig. 3 shows the Stanton number in the symmetry plane near the body flap, comparing the experimental values with computational results. The error in measured Stanton numbers is between 10 and 15% (Ref. 29). The flow is laminar; hence, the physical modeling is relatively simple. But due to the separation on the body flap, the flow features are rather complex, and none of the calculations is able to predict correctly the Stanton number in this region. Differences among the computed results and between the calculations and the experiments are attributed to differences in solution strategy, numerical schemes, different grid refinement levels in the body flap region, different convergence levels, and whether the wake and base flow were calculated.

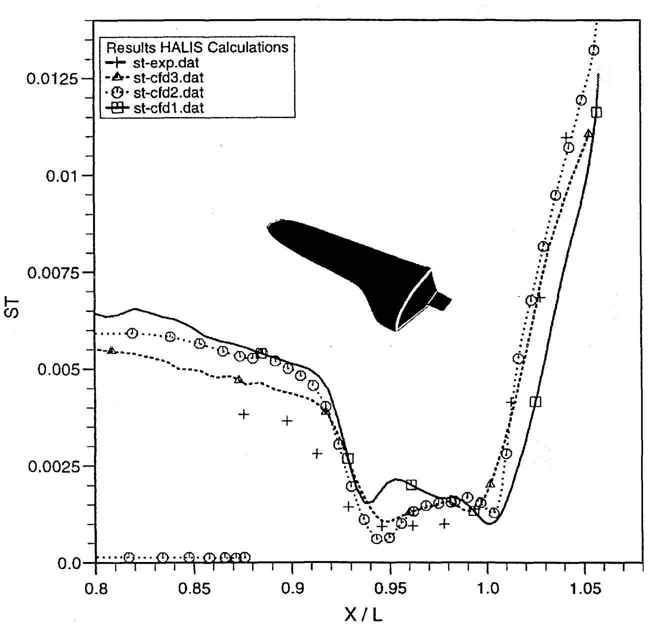


Fig. 3 Axial distribution of Stanton number in the symmetry plane in the region of the body flap, HALIS test case: $M_\infty = 9.8$, $\alpha = 40$ deg, and body-flap angle = 15 deg.

On the level of the aerodynamic coefficients, agreement existed for the C_L and C_D for the computations, but the computations differ by up to 6% from the experimental values. The calculated pitch-up moments show large variations and are up to 10% higher than the measured value.

This and other synthesis figures together with a table of aerodynamic coefficients are saved in the HAEDB, together with a report written by experts discussing the synthesis documents. Once a new computation is available (or when a new workshop is organized), it can be included in the HAEDB, and the synthesis figures and table can be updated to include the new results. Because updating the data on the Web is done easily, keeping an electronic database current is much simpler than with the conventional databases of the past.

C. Database of the Future

In Europe, the ERCOFTAC recently created an industrial advisory committee to define the flow taxonomy for flows in different industrial sectors, including aerospace. The logical next step is to define for each flow feature a representative test case with reliable experimental data. The data for these test cases could form the building blocks to construct database systems similar to the HAEDB, which can be used by people from academia and industry to validate their codes. At the same time, these database systems can provide users of CFD with the knowledge on which physical models together with their associate constants give the best prediction for a given flow feature.

Such a database system should be accessible on the WWW, with restrictive access for sensitive data; include both experimental and computational data together with reports and expert comments; offer data manipulation tools to generate synthesis plots; and allow users to add comments to results. People from academia and industry who would like to validate their code can ask for the experimental data, together with the geometrical description and physical parameters of the test case. They can run the validation exercises locally, and once they are satisfied with their results, they can send their computed results to the database for inclusion. Synthesis documents, which compare computations with experiments, can then be produced/updated. Expert comments can be attached to these synthesis documents to explain differences between computations and between experiment and computations.

A substantial investment is needed to develop such a system, including the workers to assess the quality of each experiment, hardware and software, and costs to run validation test cases. Money is also needed to operate it, including the database manager costs and the hardware replacement and/or upgrade and software maintenance. Each industry and research group developing and using CFD codes should have an interest in using such a future CFD validation database; however, none of them has the means to invest alone in such a system. A joint effort from industry and research, with funding on a governmental (or European) level, is needed to build and maintain such a system. But before building such system, agreements must be reached on data exchange standards for experiment and computations.

IX. Conclusions

Validating a simulation and assessing the credibility of a prediction based on it is a major undertaking and a difficult task, even if the estimate needs to be qualitative only. Validation is carried out at various stages of configuration complexity and by various groups for distinctly different intended uses. We have argued that an approach to establishing credibility in a simulation is via a systematic process of validation. An essential step in this process is defining a taxonomy of generic flow classes. The next step is to collect for each of these generic flow classes a well-documented test case with reliable data. These data, together with computational data, can be saved in dynamic electronic databases accessible using the WWW. The CFD community can access these databases, download data, provide new data, and comment on the results. These dynamic electronic databases also provide the CFD user with the information on which physical model to use for a given flow class to obtain a result with a sufficient level of credibility and, as such, form a handbook on how to use CFD. We have illustrated this concept with a few examples that exist already, but we believe that this is just the beginning, and soon much development and many advances will follow rapidly.

Acknowledgment

We are grateful to A. Heddergott for clarification of data reported in Ref. 11.

References

[1] Melnik, R. E., Siclari, M. J., Marconi, F., and Barber, T., "An Overview of a Recent Industry Effort at CFD Code Certification," AIAA Paper 95-2229, June 1995.

[2] Marvin, J. G., and Holst, T. L., "CFD Validation for Aerodynamic Flows—Challenge for the 90's," AIAA Paper 90-2995, Aug. 1990.

[3] Marvin, J. G., "Dryden Lectureship in Research, A Perspective on CFD Validation," AIAA Paper 93-0002, Jan. 1993.

[4] Bradley, R. G., "CFD Validation Philosophy," *Validation of Computational Fluid Dynamics*, AGARD CP 437, Vol. 1, 1988, pp. 1.1–1.6.

[5] *Validation of Computational Fluid Dynamics*, AGARD CP 437, Vol. 1, 1988.

[6] Mehta, U. B., "Guide to Credible Computer Simulations of Fluid Flows," *Journal of Propulsion and Power*, Vol. 12, No. 5, 1996, pp. 940–948.

[7] Mehta, U. B., "Computational Requirements for Hypersonic Flight Performance Estimation," *Journal of Spacecraft and Rockets*, Vol. 27, No. 2, 1990, pp. 103–112.

[8] Melnik, R. E., Siclari, M. J., Barber, T., and Verhoff, A., "A Process for Industry Certification of Physical Simulation Codes," AIAA Paper 94-2235, June 1994.

[9] Oberkampf, W. L., "A Proposed Framework for Computational Fluid Dynamics Code Calibration/Validation," AIAA Paper 94-2540, June 1994.

[10] Leyland, P., Vos, J. B., Van Kemenade, V., and Ytterström, A., "NSMB: A Modular Navier–Stokes Multiblock Code for CFD," AIAA Paper 95-0568, Jan. 1995.

[11] "Experimental Data Base for Computer Program Assessment," AGARD AR 138, Paper B3, May 1979.

[12] "Guide to the Expression of Uncertainty in Measurements," International Organization for Standardization, Geneva, Jan. 1993.

[13] "Assessment of Windtunnel Data Uncertainty," AGARD-AR-304, April 1994.

[14] "Assessment of Windtunnel Data Uncertainty," AIAA Standard S-071-1995, AIAA, Washington, DC, March 1995.

[15] Oberkampf, W. L., and Aeschliman, D. P., "Joint Computational/Experimental Aerodynamics Research on a Hypersonic Vehicle: Part 1, Experimental Results," *AIAA Journal*, Vol. 30, No. 8, 1992, pp. 2000–2009.

[16] Moffat, R. J., "Describing the Uncertainties in Experimental Results," *Experimental Thermal and Fluid Science*, Vol. 1, 1988, pp. 3–17.

[17] Rizzi, A., and Viviand, H. (eds.), *Numerical Methods for the Computation of Inviscid Transonic Flows with Shock Waves*, Vol. 3, Notes on Numerical Fluid Mechanics, Vieweg, Braunschweig, Germany, 1981.

[18] Bristeau, M. O., Glowinski, R., Periaux, J., and Viviand, H. (eds.), *Numerical Simulation of Compressible Navier–Stokes Flows*, Vol. 18, Notes on Numerical Fluid Mechanics, Vieweg, Braunschweig, Germany, 1987.

[19] Kordulla, W. (ed.), *Numerical Simulation of the Transonic DFVLR-F5 Wing Experiment*, Vol. 22, Notes on Numerical Fluid Mechanics, Vieweg, Braunschweig, Germany, 1988.

[20] Dervieux, A., Van Leer, B., Periaux, J., and Rizzi, A. (eds.), *Numerical Solution of Compressible Euler Flows*, Vol. 26, Notes on Numerical Fluid Mechanics, Vieweg, Braunschweig, Germany, 1989.

[21] Holst, T. L., "Viscous Transonic Airfoil Workshop Compendium of Results," *Journal of Aircraft*, Vol. 25, No 12, 1988, pp. 1073–1087.

[22] Kline, S. J., Cantwell, B. J., and Lilley, G. M. (eds.), *1980-81 AFOSR-HTTM-Stanford Conference on Complex Turbulent Flows*, Vols. 1–3, Mechanical Engineering Dept., Stanford Univ., Stanford, CA, 1981.

[23] Abgrall, R., Desideri, J.-A., Glowinski, R., Mallet, M., and Periaux, J. (eds.), *Proceedings of INRIA-GAMNI/SMAI Workshops on Hypersonic Flows for Re-Entry Problems, Parts I and II*, Springer–Verlag, Berlin, 1992.

[24] Bonnin, J. C., and Rodi, W.,"Progress Report on the SCIENCE Project 'Data Validation and Comparison in Fluid Mechanics,' " ERCOFTAC Bulletin 22, Sept. 1994.

[25] Bonnin, J. C., Buchal, T., and Rodi, W., "ERCOFTAC Workshop on Data Bases and Testing of Calculation Methods for Turbulent Flows," ERCOFTAC Bulletin 28, March 1996.

[26] Fleming, J. L., Simpson, R. L., Cowling, J. E., and Devenport, W. J., "An Experimental Study of a Turbulent Wingbody Junction and Wake Flow," *Experiments in Fluids*, Vol. 14, 1993, pp. 366–379.

[27] Jones, W. P., and Launder, B. E., "The Prediction of Laminarization with a Two-Equation Model of Turbulence," *International Journal of Heat and Mass Transfer*, Vol. 15, No. 2, 1972, pp. 301–314.

[28] Hunt, J. J. (ed.), *Second European Symposium on Aerothermodynamics for Space Vehicles*, European Space Agency, ESA SP-367, Noordwijk, The Netherlands, 1995.

[29] Guillemet, V., Stojanowski, M., and Bugeau, A., "Heat Flux and Pressure Measurements on Halis Models in S4 Modane Wind Tunnel," *Proceedings of the 4th European High Velocity Data Base Workshop*, edited by J.-A. Desideri and J. Periaux, Institut National de Recherche en Information et en Automatique, INRIA Rept. 94-1021, Valbonne Cedex, France, 1994.

W. Oberkampf
Associate Editor