Australian Government
Department of Defence
Defence Science and
Technology Organisation

DSTO

Probabilistic Methods for the Quantification of Uncertainty and Error in Computational Fluid Dynamics Simulations

John Faragher

DSTO-TR-1633

Australian Government
Department of Defence
Defence Science and
Technology Organisation

Probabilistic Methods for the Quantification of Uncertainty and Error in Computational Fluid Dynamics Simulations

John Faragher

Air Vehicles Division
Platforms Sciences Laboratory

DSTO-TR-1633

ABSTRACT

The life of engine components is determined by a combination of the material properties and the applied stresses and temperatures. As a consequence of variability in these parameters, the component life is not fixed (deterministic) but stochastic (random) and may be characterised by a probability density function (PDF). In order to reduce the cost of ownership of ADF aircraft these PDFs need to be determined as accurately as possible. Probabilistic techniques offer significant potential for accurate and realistic estimates of component lives by quantifying stochastic elements of an analysis rather than introducing excessive conservatism to allow for them. This report examines the feasibility of using a probabilistic approach for modelling the component temperatures in an engine using CFD (Computational Fluid Dynamics).

Published by

DSTO Platforms Sciences Laboratory
506 Lorimer St
Fishermans Bend, Victoria 3207 Australia

Telephone: (03) 9626 7000
Fax: (03) 9626 7999

AR-013-229
October 2004

APPROVED FOR PUBLIC RELEASE

Probabilistic Methods for the Quantification of Uncertainty and Error in Computational Fluid Dynamics Simulations

Executive Summary

DSTO has developed a risk and reliability management capability for mechanical components in propulsion systems. This includes an assessment of relative airworthiness risks resulting from modifications to life limits as a consequence of operational, logistic and repair requirements. It will expand the understanding of the fundamental mechanisms and processes responsible for the statistical distribution of lives in aircraft dynamic components. The overall purpose of this capability is to provide technical airworthiness advice to the RAAF on those critical factors that underpin the promulgation of operational lives of propulsion system components, and on the risks associated with varying these life limits.

The life of engine components is determined by a combination of the material properties and the applied stresses and temperatures. As a consequence of variability in these parameters, the component life is not fixed (deterministic) but stochastic (random) and may be characterised by a probability density function (PDF). In order to reduce the cost of ownership of ADF aircraft these PDFs need to be determined as accurately as possible. Probabilistic techniques offer significant potential for accurate and realistic estimates of component lives by quantifying stochastic elements of an analysis rather than introducing excessive conservatism to allow for them. The objective is to devise a methodology that brings together probabilistic and deterministic approaches in a manner that best supports improved component life estimates. This will reduce under-utilisation and cost of spare parts, and increase safety.

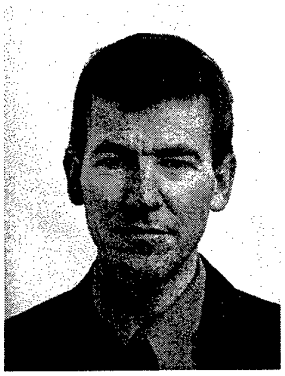
This report examines the feasibility of using a probabilistic approach for modelling the component temperatures in an engine using CFD (Computational Fluid Dynamics). Currently, CFD analysis for engine temperature uses deterministic inputs. Small variations in these inputs may have a significant effect on the resultant component temperatures and in turn the predicted life. This is especially important for creep life assessment when a change of 20°C can halve the creep life.

The CFD discipline is less mature than the linear finite element stress analysis discipline because CFD requires much greater computer power to solve problems of practical engineering interest. For this reason probabilistic methods have been too computationally expensive to be widely adopted in CFD. Computer power has now increased to the point where probabilistic CFD has become a very active area of research.

This report untangles some of the confusion and debate which currently surround the definitions of "error", "uncertainty" and "variability" in computational simulation, and the three distinct processes of verification, validation and uncertainty analysis. It critically reviews a range of probabilistic and non-probabilistic methods which may be used to propagate the uncertainty from the input variables, through the model, to the output variables. Finally it makes recommendations for the application of the probabilistic methods to CFD.

Authors

John Faragher
Air Vehicles Division



John Faragher obtained a B.E.(Hons) from the University of Melbourne in 1987. He commenced work in the Propulsion Research Area at the Aeronautical and Maritime Research Laboratory in 1988. He worked in the areas of Engine Performance and Engine Mechanical Integrity including research into engine component lifing. He was awarded a cadetship to study full-time at the University of Melbourne from 1992 to 1995 and was awarded a Ph.D. (Engineering). Since 1995 he has continued to work in the Propulsion Research Area at the Aeronautical and Maritime Research Laboratory (now called Platform Sciences Laboratory). In addition to the work described in this report, he has been developing computer models for research into gas turbine engine temperature modelling and heat transfer problems.

Contents

1. Introduction

DSTO has developed a risk and reliability management capability for mechanical components in propulsion systems. This includes an assessment of relative airworthiness risks resulting from modifications to life limits as a consequence of operational, logistic and repair requirements. It will continue to expand the understanding of the fundamental mechanisms and processes responsible for the statistical distribution of lives in aircraft dynamic components. The overall purpose of this work is to provide technical airworthiness advice to the RAAF on those critical factors that underpin the promulgation of operational lives of propulsion system components, and on the risks associated with varying these life limits.

Computational simulation and analysis of engineering systems offers huge benefits when it is used wisely. In many industries it has drastically reduced the time and cost involved in the traditional engineering design cycle of: design, build, test, re-design, re-build, re-test, etc.. It has also been used to simulate experiments that are too difficult or dangerous to perform, such as accidents involving nuclear weapons and space vehicle atmospheric entry. Unfortunately, the advent of "user-friendly" commercial software and cheap computer hardware has placed inexperienced analysts in the dangerously tempting position of being able to perform sophisticated simulations and produce impressive three-dimensional, animated visualisations without knowing how accurate the results are. The need for appropriate training and experience has never been greater. Various professional organisations are addressing this issue by producing "guidelines for best practice". This report explains some of the complexities involved in defining and quantifying the sources of inaccuracy and uncertainty, and their effect on the results of computational fluid dynamics (CFD).

Traditional engineering design and analysis methods are deterministic. Mean values of the design variables and physical constants are used to calculate single-point estimates of the system behaviour. These single-point estimates of behaviour enable engineers either to analyse an existing system or to optimise the design of a new system. The mean values of the design variables include data defining the geometric, material, structural, thermal and aerodynamic properties of the design. In reality the data are not known exactly because they contain uncertainties, such as limited accuracy in measurements and tolerances in manufacturing. This forces the design engineer to err on the side of conservatism and introduce "safety factors" which may be very large.

The life of engine components is determined by a combination of the material properties and the applied stresses and temperatures. As a consequence of variability in these parameters, the component life is not fixed (deterministic) but stochastic (random) and may be characterised by a probability density function (PDF). In order to reduce the cost of ownership of ADF aircraft these PDFs need to be determined as accurately as possible. Probabilistic techniques offer significant potential for accurate and realistic estimates of component lives by quantifying stochastic elements of an

analysis rather than introducing excessive conservatism to allow for them. The objective is to devise a methodology that brings together probabilistic and deterministic approaches in a manner that best supports improved component life estimates. This will reduce under-utilisation and cost of spare parts, and increase safety.

Probabilistic methods have been used in aircraft engine component structural design and fatigue life prediction for decades (Yang & Chen 1985, Fox 1994, Kappas 2002). Similarly, aircraft operators, particularly the U.S. Air Force, have been using non-deterministic methods to address ageing-aircraft fatigue issues for decades (Lincoln 1985, Tong 2001). On the other hand, the application of probabilistic methods to airframe structural design and aerodynamics by aircraft manufacturers is only just beginning (Zang et al. 2002). The CFD discipline is less mature than the linear finite element stress analysis discipline because CFD requires much greater computer power to solve problems of practical engineering interest. For this reason probabilistic methods have been too computationally expensive to be widely adopted in CFD. Now, computer power has increased to the point where probabilistic CFD is "coming of age" (Walters and Huyse 2002).

This report examines the feasibility of using a probabilistic approach for modelling the component temperatures in an engine using CFD. Small variations in the inputs may have a significant effect on the resultant component temperatures and in turn the predicted life. This is especially important for creep life assessment when a change of 20°C can halve the creep life.

Section 2 of this report untangles some of the confusion and debate which currently surround the definitions of "error", "uncertainty" and "variability" in computational simulation. Section 3 provides an overview of the steps involved in a CFD simulation. Section 4 classifies and discusses the sources of error in CFD. Section 5 clarifies the three distinct steps which are necessary to estimate the magnitude of the uncertainty from all sources in a CFD simulation. Section 6 critically examines a range of probabilistic and non-probabilistic methods which may be used to propagate the uncertainty from the input variables, through the model, to the output variables. Section 7 presents the conclusions and recommendations.

2. Overview of CFD Uncertainty, Variability and Error

2.1 Definition of Error, Uncertainty and Variability

The accuracy of a CFD simulation is adversely affected by error and uncertainty. The American Institute of Aeronautics and Astronautics (AIAA) "Guide for the Verification and Validation of CFD Simulations" (AIAA 1998) suggests the following definitions of error and uncertainty:

Error: A recognisable deficiency that is not due to lack of knowledge.

Uncertainty: A potential deficiency that is due to lack of knowledge.

These two definitions imply that error is deterministic in nature and uncertainty is stochastic or non-deterministic in nature. In simple terms, this means that if a simulation containing an error is run twice it will produce the same answer (there are exceptions). On the other hand, if a simulation containing an uncertainty is run twice it will produce a different answer each time - assuming, of course, that the uncertainty is modelled as an uncertain quantity and not replaced by a mean or nominal value. Errors are not due to lack of knowledge and can be identified and corrected, given enough time and money. In most cases errors will be allowed to remain in a simulation if they are estimated to be within reasonable limits and there is not enough time and money available to eliminate them.

Both errors and uncertainties can be further broken down into two types.

Errors can be classified as either "acknowledged" or "unacknowledged":

- Acknowledged errors are those which have been identified in a simulation but the analyst decides not to remove them because the results are adequate for the job requirements and budget (examples include round-off error and discretization error, and convergence error in an iterative numerical scheme).

- Unacknowledged errors in a simulation are those the analyst is unaware of but they are recognisable. They have no foolproof procedures for finding them but they may be detected by redundant procedures and double-checking (examples are computer programming errors, or usage errors, including mistakes and blunders) .

Uncertainty can be classified as either "aleatory" or "epistemic". The risk analysis community was the first to make this distinction in practical applications.

- Aleatory uncertainty is also called "variability", "stochastic uncertainty", "inherent uncertainty" or "irreducible uncertainty". It is the physical variation present in the system being analysed or its environment. It is not due to a lack of knowledge and cannot be reduced.

- Epistemic uncertainty is also called "reducible uncertainty" or simply "uncertainty". This is what the AIAA Guide (AIAA 1998) defines as "uncertainty", i.e. a potential deficiency that is due to a lack of knowledge. Epistemic uncertainty can be reduced by an increase in knowledge about the system being analysed. The other kind of uncertainty (aleatory) is not mentioned in the AIAA Guidelines.

The distinction between "variability", "uncertainty", and "error" is important not only when assessing how each contributes to an estimate of total modelling and simulation uncertainty, but also when determining how each should be represented mathematically and propagated through the mathematical model. The propagation of variabilities, uncertainties and errors from their sources to the calculated CFD results is called "uncertainty analysis". This area is currently the subject of fierce debate in the literature. Oberkampf and his colleagues at Sandia National Laboratory (Oberkampf et al. 2000) believe that the common practice of representing and propagating aleatory uncertainty (variability) by using probability density functions is appropriate, but they disagree with representing epistemic uncertainty in the same way. They argue that because it is caused by a lack of knowledge it should be represented using such mathematical theories as "evidence theory" (Guan and Bell 1991), "possibility theory" (de Cooman et al. 1995), "fuzzy set theory" (Cox 1999) or "imprecise probability theory" (Kozine 1999).

Uncertainty analysis methods are being adopted because the degree of conservatism in traditional engineering design is sometimes excessive. The uncertainty analysis methods allow engineers designing new components to apply more rigour to the determination of safety factors, and owners of older equipment to assess the safe lifetime of their equipment and to quantify the risks involved in any life extension. Uncertainty analysis methods assign a probability distribution to each uncertain variable around its mean value, and then propagate this uncertainty through the mathematical model to the output. The effect of the uncertainties on the predicted performance and endurance of the component can then be quantified.

For example, a structural analysis of the Space Shuttle Main Engine performed at NASA Glenn using probabilistic methods found that uncertainties in the geometry data had a statistically significant effect on the behaviour of the components but, in contrast, uncertainties in the material property data had a statistically insignificant effect (Nagpal et al. 1987).

In the literature, there is very little consistency in the definitions of "variability", "uncertainty", and "error". Many authors do not clearly define what they mean when they use these terms. Or their definitions contradict different authors within the same discipline, and sometimes even the same journal. Some books in the field of reliability, uncertainty and risk analysis do not address the issue of numerical solution error. This can be particularly detrimental to uncertainty estimation when the mathematical models involve partial differential equations (PDEs) (Oberkampf et al. 2000).

Further confusion is added when authors refer to solution error as "numerical uncertainty" when they mean "numerical inaccuracy" or "numerical error". As mentioned above the distinction between error and uncertainty is important. Numerical solution error results from the discrete approximation of a continuous

system. It does not result from a lack of knowledge. The approximation is a conscious decision made by the analyst. Therefore it is a form of error, not of uncertainty.

The lack of agreement in this area is illustrated by the following excerpt from the AIAA guidelines (AIAA 1998):

"It is important to emphasize that this document presents guidelines for V&V of CFD simulations, not standards. The AIAA CFD Committee on Standards unanimously believes that the state of the art in CFD has not developed to the point where standards can be written."

2.2 Verification and Validation

It is important to understand the distinctions between a computer code, a simulation or analysis, and a mathematical model. A mathematical model and a computer code are tools which are used to perform an analysis or simulation. The tools and processes involved in a computer simulation without uncertainty analysis are shown in Figure 1. A mathematical model is an approximate representation of the behaviour of a real physical system. It is based on theoretical hypotheses and empirical correlations. A computer code uses numerical techniques to produce an approximate solution of a mathematical model. It solves only a discrete approximation to the continuous equations. The code requires input data which may not be precisely known. The computer code is used to perform a CFD simulation or analysis which yields values used by the engineer to design a new system or analyse an existing one.

An inexperienced analyst could easily be misled into thinking that verification and validation (V&V) is a tautology describing a single process; both because they are so often talked about together and because the two words are listed as synonyms in a thesaurus of common english usage. In the technical language of computer simulation they are widely accepted as describing two quite separate and distinct processes. This is illustrated in Figure 1 and can be defined as follows:

Verification (are we solving the equations correctly?):

- estimates the magnitude of the error in the computational implementation of the mathematical model.
- compares the numerical methods used in the code to exact analytical results.
- tests for computer programming errors.

Validation (are we solving the correct equations?):

- estimates the magnitude of the difference between the results of the computational simulation and physical reality.
- compares the computed results with experimental results.

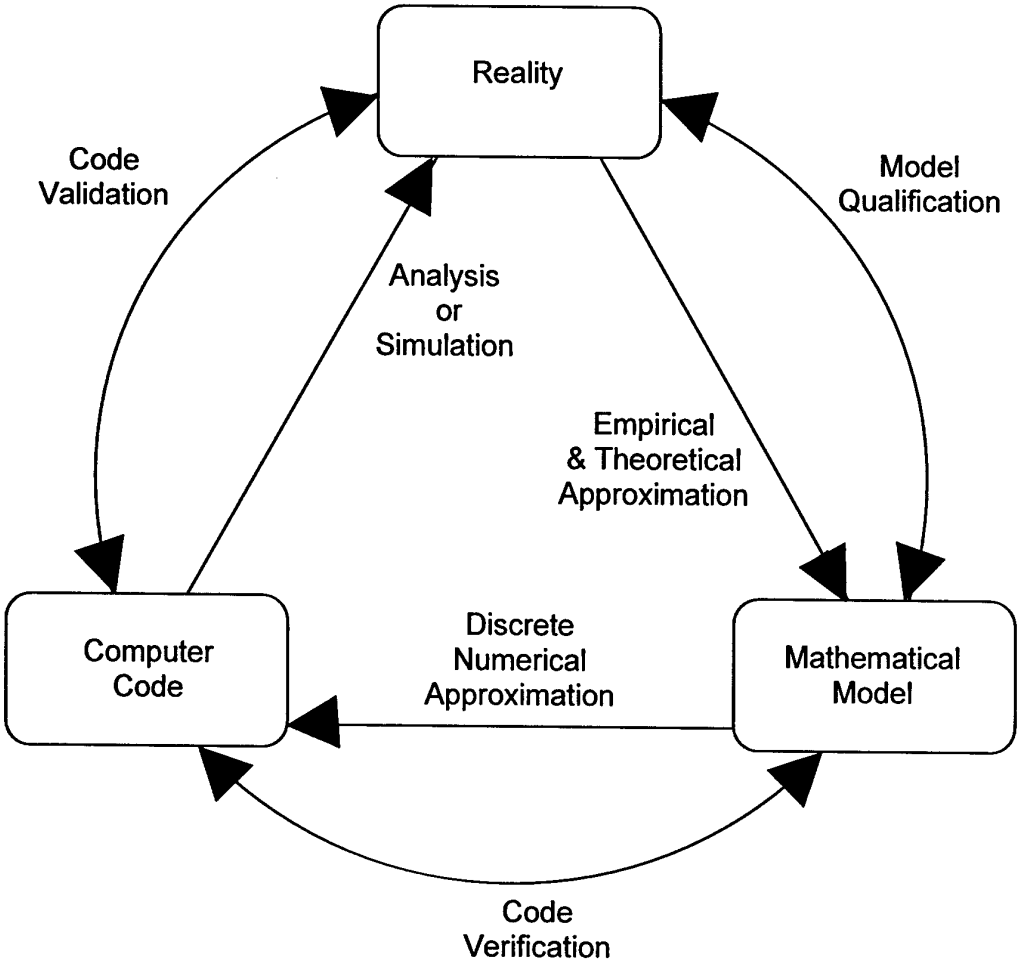


Figure 1. The Phases of Computer Simulation Without Uncertainty Analysis

The definition of "prediction" in the AIAA guidelines is very interesting and would surprise some people because it explicitly does not include calculating results for conditions for which the CFD code has been validated. This is not "prediction" but "interpolation" within the experimental data because the validation process is, by definition, a successful comparison of the code with experimental data. "Prediction" is defined as using a CFD model to foretell the state of a physical system under conditions for which the CFD model has **not** been validated. An estimation of the uncertainty in a CFD simulation must therefore include an estimate of how far the prediction state is from the validation state, and how much uncertainty this introduces to the results.

It is important to note that the process of validation typically does not include uncertainty analysis. Thus the uncertainty in the inputs is not quantified and not propagated through the model to the outputs. The input quantities are taken from the

experimental data as mean or nominal values and are treated as if they are true values. This neglects the uncertainty in the experimental measurement of the inputs. It also neglect deficiencies in the experimental data, such as quantities that must be specified as inputs to the CFD code that were not measured in the experiment or not measured with sufficient detail, for example inlet velocity profiles.

There is a growing recognition in the fluid dynamics community that data from older experiments, which were not designed for the validation of CFD codes, are often not suitable for this purpose. This was emphasised by the NATO AGARD Fluid Dynamics Panel in 1994 when it published a selection of test data for the validation of CFD codes (AGARD 1994) and stated:

"*The Working Group found it difficult to select reliable test cases. The inclusion of a test case within the database does not automatically guarantee good quality. The working Group takes no responsibility for the fitness or otherwise of the database information. ... For that reason, AGARD FDP would appreciate it very much if the experience with particular test cases could be reported to the Chairman of AGARD's FDP Committee on Wind Tunnel Testing Techniques.*"

The true value of a quantity is rarely known. The nearest thing to a true value is a "standard" used for calibrating experimental equipment. Neither experimental measurements nor computational predictions give the true value of a quantity. Both contain errors and uncertainties which must be estimated. There are official standards describing methods for estimating uncertainty in experimental measurements (ANSI/ASME 1985, ISO 1995, AIAA 1999). There is professional disagreement on the exact procedures for verification and validation of CFD simulations. No official standards for CFD verification and validation exist.

3. The CFD Analysis Process

In order to discuss the errors and uncertainties in a CFD simulation, it is necessary to understand the steps in the process. These are:

- Understand the Flow Problem and Formulate the Simulation Strategy
- Create the Geometry
- Generate the Mesh
- Establish the Boundary and Initial Conditions
- Run the Solver, and Monitor for Convergence
- Post-Process and Interpret the Results
- Refine the Mesh
- Perform a Sensitivity Analysis
- Document the Analysis

In further detail, these steps include:

3.1 Understand the Flow Characteristics and Formulate the Simulation Strategy

When applying CFD to flows in aerospace systems it is vital to understand the characteristics of the flow. These may include Mach number, Reynolds number, turbulence, boundary layers, adverse pressure gradients, shock waves and separation. At hypersonic Mach numbers, real gas effects may become important.

A useful first step is to formulate the simulation strategy by answering the following questions:

- what is the objective of the analysis?
- what degree of accuracy is required?
- what geometry should be included?
- what are the freestream and/or operating conditions?
- what dimensionality of the spatial model is required? (2D, axisymmetric, 3D)
- what should the flow domain look like?
- how far away should the domain boundaries be?
- what near-wall treatment will be used at walls?
- what temporal modelling is appropriate? (steady or unsteady)
- what is the nature of the viscous flow? (inviscid, laminar, turbulent)
- what turbulence model will be used? (if any)
- what solver will be used? (segregated, coupled, implicit, explicit)
- how should the gas be modelled (incompressible, compressible)?
- is heat transfer important?
- is heat conduction through solids to be modelled?
- which radiation heat transfer model is to be included? (if any)
- are chemical species or reactions to be modelled?
- will the problem fit in the memory of the computer?
- how long will the problem take to converge?

The accuracy required of CFD for engineering analysis may be divided into three levels:

1) to provide qualitative information,
2) to provide incremental quantities, and
3) to provide absolute quantities.

3.2 Create the Geometry

The geometry of the body about which or inside which the flow is to be analysed is usually created in a CAD software package. Simplifications of the geometry may be

required to allow an analysis to be performed with reasonable effort. This requires an understanding of which geometrical features are significant to the flow and which features can be neglected. The extent of the flow domain must be defined. The boundary of the flow domain may comprise solid surfaces, inlets, outlets, symmetry, periodic, axis and far-field boundaries. To approximate true infinite-extent conditions effectively, the far-field boundaries must be placed far enough from the object of interest not to interfere with the flow. For example, in lifting airfoil calculations, it is common for the far-field boundary to be a circle with a radius of 20 chord lengths. The creation of the geometry and flow domain takes into account the planned structure and topology of the mesh generation.

3.3 Generate the Mesh

Mesh generation is a vast mathematical specialty of its own. In CFD the flow domain is discretised into a mesh made up of points defining the corners of cells. In 2D, quadrilateral or triangular cells are used; and in 3D, hexahedral, tetrahedral, pyramid, or wedge cells may be used. Meshes may be single-block or multi-block structured meshes, unstructured or hybrid. In multi-block meshes subdomain boundaries may be conformal (all mesh points coincident), non-conformal or overlapping. The mesh generation involves defining the structure and topology, and then generating a mesh on that topology. Solvers using the finite-difference method require structured meshes. Structured meshes are also called grids because they can be transformed mathematically into a uniform rectangular grid. Solvers using the finite-volume method allow unstructured meshes which have no regularity. This greatly simplifies the task of mesh generation for complex geometries and local mesh refinement.

The mesh should exhibit some minimal mesh quality as defined by measures of orthogonality (especially at the boundaries), relative mesh spacing (15% to 20% stretching is considered a maximum value), mesh skewness, etc.. Further, the maximum spacings should be consistent with the desired resolution of important features. The resolution of boundary layers requires the mesh to be clustered in the direction normal to the surface, with the spacing of the first mesh point off the wall to be well within the laminar sublayer of the boundary layer. For turbulent flows, the first point off the wall should exhibit a y+ value of less than 1.0. Quadrilateral/hexahedral cells may be more economical in some situations because they can tolerate a much larger aspect ratio than triangular/tetrahedral cells. A large aspect ratio in a triangular/tetrahedral cell will invariably affect the skewness of the cell, which is undesirable as it may impede accuracy and convergence. Poor quality meshes may produce no solution at all or, more dangerously, an inaccurate solution.

3.4 Establish the Boundary and Initial Conditions

Flow and thermal variables must be specified on the boundaries of the flow domain. In many cases, the definition of an appropriate initial solution for the solver to start from is essential to obtaining a satisfactory final solution. Material properties must be specified for all solids, fluids and chemical species included in the simulation.

3.5 Run the Solver and Monitor for Convergence

The simulation is performed with options for interactive or batch processing. The domain may need to be decomposed into an appropriate number of partitions for parallel processing. At each step, the solution is monitored to determine whether a "converged" solution has been obtained. Solver parameters, such as under-relaxation, may need to be adjusted to enable the solver to converge to a solution. If convergence is not achieved the solver may need to be restarted with different initial conditions. Alternatively, the boundary or free-stream conditions, such as a flow velocity, may need to be relaxed to achieve an intermediate solution from which a final solution can be approached by gradually increasing the boundary conditions to the required values.

3.6 Post-Process and Interpret the Results

Post-Processing involves extracting the desired flow properties (pressure, temperature, velocity, etc...) from the computed flow field. It may make use of a large range of tools to produce anything from simple plots through to colourful three-dimensional, animated graphics. This is also called "visualisation of the solution". Post-processing often involves deriving new results from calculations using the data produced by the solver. At this stage of the CFD simulation process new errors may be introduced. These errors may be caused either by the inexperience of the user, or errors in the post-processing software.

The solution is not necessarily correct just because it has converged. Engineering judgement and experience must be applied to decide whether the behaviour of the predicted flow is believable.

3.7 Refine the Mesh

If the solver is unstructured it is possible to refine the mesh locally. Solution-adaptive mesh refinement tools allow the analyst to refine and/or coarsen the mesh according to criteria based on the solution data. Gradient based refinement tools allow mesh points to be added in regions of the flow field where there are large gradients of a particular variable, such as pressure, density or temperature. These regions of the flow may be where boundary layers, shear layers or shock waves occur. By using solution-adaptive

refinement, cells are only added where they are needed in the mesh, thus enabling the important features of the flow field to be resolved more accurately without wasting computational resources.

3.8 Perform a Sensitivity Analysis

The sensitivity of the computed results may be examined with respect to such variables as:

- dimensionality
- flow conditions
- initial conditions
- boundary conditions
- algorithms
- mesh topology and density
- turbulence model
- chemistry model
- radiation flux model
- artificial viscosity
- computer system

3.9 Document the Analysis

Documenting the findings of an analysis involves describing each of these steps in the process with enough detail for another analyst to be able to repeat it.

3.10 A Final Note on the CFD Simulation Process

This section has discussed the steps in the CFD simulation process in order to provide the context for defining, in the next section, the errors and uncertainties that may arise. It has also highlighted the complexity of the process and the consequent need for expertise and experience in the analyst. This is well expressed by the "Best Practice Guidelines For Marine Applications of CFD" produced by the MARNET-CFD Thematic Network supported by the European Commission in the area of Quality and Trust in Industrial CFD (MARNET-CFD 2000):

"*Unlike linear finite element stress analysis, CFD still requires expertly trained users for good results. In situations where non-experienced users have to be used, some restriction on their freedom to adjust critical parameters might be advisable, and they should be limited to simulations of routine types.*"

4. Sources of Error in CFD

4.1 Taxonomy or Classification of Errors and Uncertainties

A taxonomy or taxonomic classification consists of a set of mutually exclusive sets. This ensures that each item can appear in only one class or set and eliminates confusion resulting from counting the same item twice. A taxonomy of errors in CFD has been created by one author from the Validation and Uncertainty Estimation Department of Sandia National Laboratories (Oberkampf et al. 1995) and endorsed by the most frequently cited textbook of V&V (Roache 1998). It comprises four categories:

1. Physical Modelling Errors
2. Discretization Errors
3. Programming Errors
4. Computer Round-Off Errors

This taxonomy applies to the tools used for computer simulation (the mathematical model and the computer code) and therefore does not include user errors caused by poor practices which may result from inadequate training, or from mistakes. It also does not include the error called "iterative convergence error" which occurs either when the code cannot achieve satisfactory convergence to a solution, or when the user stops the code because it is taking longer to converge than the time available.

4.2 Physical Modelling Errors

This includes simplifications made to the analysis for reasons of economy and based on the assumption that their effect on the results will be small enough to be acceptable. For example:

- using a steady solver to simulate unsteady flow
- using RANS (Reynolds-averaged Navier-Stokes) instead of DNS (Direct Numerical Simulation)
- using wall functions instead of a fine mesh near the walls
- placing far-field boundaries a finite distance from the region of interest
- modelling a compressible flow as incompressible
- modelling a viscous flow as inviscid
- modelling temperature dependent properties as constant
- using the Boussinesq approximation for natural convection
- modelling a mixture of gas species as a single gas, e.g. modelling engine exhaust gas as air
- using a simple radiation model

4.3 Discretization Errors

Discretization errors can be evaluated by grid refinement and time step refinement studies. As the grid size and time step are reduced the discretisation error should asymptotically approach zero. If it does not there are programming errors in the code. If it does, then the Richardson Extrapolation technique can be used to estimate what the error would be if the grid size and time step were actually infinitesimal (Richardson 1927, Roache 1998).

There is on-going debate about the pros and cons of structured 2D quadrilateral or 3D hexahedral meshes, as opposed to un-structured 2D triangular or 3D tetrahedral meshes. When the geometry is complex, a triangular/tetrahedral mesh can often be created with far fewer cells than the equivalent mesh consisting of quadrilateral/hexahedral cells. This greatly reduces the computational expense of the solution.

On the other hand, discretisation error is minimized when the flow is aligned with the mesh. It is clear that with a 2D triangular or 3D tetrahedral mesh the flow can never be aligned with the cell edges. With a 2D quadrilateral or 3D hexahedral mesh, the flow might be aligned with the cell edges for simple flows and in the layers adjacent to boundary surfaces. Discretisation error comprises two different types of error. The first type is the spatial and temporal discretisation of the flow domain. The second type, called "truncation errors", are unavoidable in practice because the computer can solve only a discrete approximation to the partial differential fluid flow equations. A mathematically exact representation of each derivative would require an infinite number of terms in each Taylor's series. If the lowest-order term that is truncated (neglected) is first-order (e.g. involves Δx) then the method is called "first-order accurate". If the first-order term is not truncated and the lowest-order term that is truncated is second-order (e.g. involves Δx^2) then the method is called "second-order accurate".

Most commercial CFD codes offer both first-order accurate and second-order accurate methods. Results calculated using a first-order accurate method are unacceptable for publication in most journals; but for complex flow problems, particularly those involving compressible flow, it may be impossible to make a second-order accurate method converge to a solution. The calculation will diverge or become unstable when the numerical errors are amplified at each step in the calculation. It may be necessary to use a first order scheme at the start of a calculation as it is likely to be more robust; but as convergence is approached a second order scheme should be used, if possible.

Another way of looking at the truncation error is useful in understanding the behaviour of the numerical solution of the flow equations. It can be shown that the exact (with zero round-off errors) numerical solution of a discretised partial differential equation (PDE), while being an approximate solution to that PDE, is the exact solution

of a slightly different PDE, which is called the "modified equation". Typically this PDE, instead of having zero on the RHS, has the terms of the truncation error on the RHS.

If the leading terms of the truncation error on the RHS of the "modified equation" are even-order derivatives, then the behaviour of the solution is dominated by "numerical dissipation" or "artificial viscosity". This artificial viscosity occurs in the first order upwind scheme, which is particularly popular in finite-volume CFD methods. Numerical dissipation is obviously not a real physical phenomenon, but its effect on a flow calculation is analogous to that of increasing the real viscosity coefficient.

On the other hand, if the leading terms of the truncation error on the RHS of the "modified equation" are odd-order derivatives, then the behaviour of the solution is dominated by "numerical dispersion" which produces oscillations in the solution. This is a characteristic wavy pattern with a wavelength of two cell sizes in the neighbourhood of steep gradients. These so-called "wiggles" are caused by dispersion errors, i.e. waves with different wave lengths are not transported with the same speed.

Artificial viscosity reduces the accuracy of a solution; but it improves the stability of the solution. Therefore it is not always seen as a bad thing. It may prevent a calculation from diverging. For some flow calculations, such as compressible flows containing shock waves, it is common practice to add extra artificial viscosity in order to achieve a solution. It raises a question that has been debated in the CFD community for several decades: is a solution using this technique better than no solution at all? The answer varies depending on the flow being considered, but the consensus is that this technique, in the hands of experts, has produced some very useful results.

The combined effect of numerical dispersion and numerical dissipation is often called "numerical diffusion", "discretization error". or "numerical error". A consistent numerical method will approach the continuum representation of the equations and zero discretization error as the number of mesh points increases to infinity and the size of the mesh spacing tends to zero. As the mesh is refined, the solution should become less sensitive to the mesh spacing and approach the continuum solution. This is called "grid convergence". These are the errors that are addressed by a "grid convergence study" or "grid refinement study". The level of discretization error depends on mesh quality. The mesh should be generated with consideration of such qualities as resolution, density, aspect ratio, stretching and orthogonality (as mentioned above in the section on "Mesh generation").

This type of error arises in all numerical methods. It is related to the approximate representation of a parameter which varies continuously in space by some polynomial function for the variation across a mesh cell. In first order schemes, for example, the parameter is taken as constant across each cell.

4.4 Round-Off Errors

In addition to truncation errors the numerical solution of the equations is affected by round-off errors. These errors arise because any computer only stores each number to a certain precision. Every time a new number is calculated it is rounded-off to that precision. Round-off errors mean that the computer cannot produce the exact solution of the discrete equations. Round-off errors can be reduced by using "double precision" options to represent numbers using 64 bits instead of 32 bits. Round-off errors are usually insignificant, but do occasionally cause major inaccuracy or prevent convergence. This can be tested by comparing calculations using 32 bits number storage with those using 64 bit number storage.

4.5 Programming Errors

Programming errors are simply mistakes. They come under the classification of unacknowledged errors described above. They may be detected by redundant procedures and double-checking or in the process of verification.

5. Sources of Uncertainty in CFD

It is now necessary to introduce a second and complementary classification of uncertainty in a CFD analysis. It divides the uncertainty into two parts. These are:

1. Uncertainty about how well the mathematical model represents the true behaviour of the real physical system. This type of uncertainty is called "model form uncertainty", "structural uncertainty", "nonparametric uncertainty" or "unmodelled dynamics". This type of uncertainty is very difficult to characterise in terms of probability density functions. In CFD, the model with the greatest amount of this type of uncertainty is the turbulence model because the phenomenon of turbulence is not fully understood.

2. Uncertainty that arises because the precise data needed for the analysis is not available. This type of uncertainty is sometimes called "parameter uncertainty", "parametric uncertainty" or "parameter variability". This class of uncertainty results from both a lack of knowledge and an inherent variability in the parameters (i.e. aleatory and epistemic uncertainty). Examples of this are: uncertainties in the precise geometry of manufactured components; and uncertain data that need to be specified as boundary conditions and material properties.

Dividing the uncertainty in this way, and using the error classifications from section 4 above, allows us to clarify the three distinct processes required to quantify the error and uncertainty from all sources in a CFD simulation as follows:

1. Verification estimates the magnitude of:
   - discretisation error, plus
   - round-off error, plus
   - programming error

2. Validation estimates the magnitude of:
   - physical modelling error, plus
   - model form uncertainty

3. Uncertainty analysis estimates the magnitude of:
   - uncertainty in the output caused by parameter uncertainty and variability

The methods used to achieve this third step are discussed in section 6.

6. CFD Uncertainty Analysis Methods

"Uncertainty analysis" methods estimate the uncertainty in the output of a CFD code from the uncertainty in the inputs. For this reason they are also called "uncertainty propagation" methods. There are three classes of uncertainty analysis methods corresponding to three different ways of characterising the uncertainty in the input variables, which are shown in Figure 3.

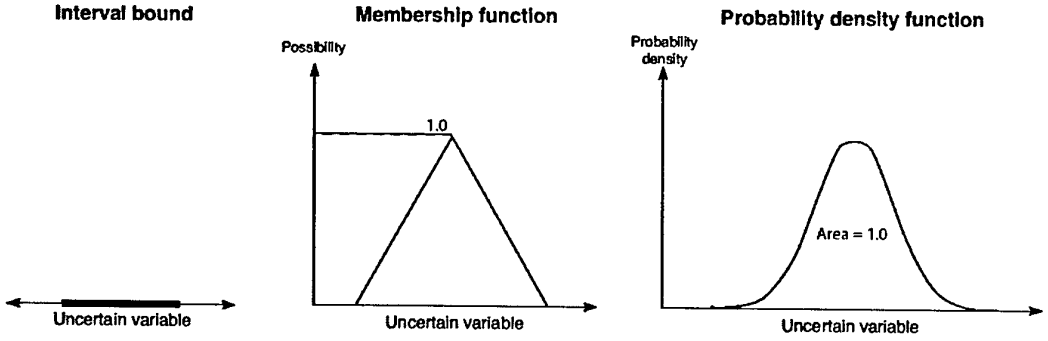


Figure 3: three different ways of characterising the uncertainty in the input variables

1. The simplest way of characterising the uncertainty in the input variables is by specifying upper and lower limits called "interval bounds". When this is the only information available it can be propagated through the CFD code to the output using "interval analysis".

2. The definition of "membership functions" requires more information. This information can be propagated through the CFD code using fuzzy logic methods.

3. If there is enough information about the input variables to define probability density functions, then "probabilistic" uncertainty analysis methods can be used. These methods are also called "stochastic" methods.

Some of the uncertainty analysis methods which have been applied to CFD simulations are described below.

6.1 Non-Probabilistic Methods

6.1.1 Interval Analysis

In interval analysis the value of a variable is replaced by a pair of numbers representing the maximum and minimum values that the variable is expected to take. Interval arithmetic rules are then used to perform mathematical operations with the interval numbers. Some Fortran 95 compilers now include these operations. While interval analysis has been applied successfully to small problems, it has not found practical application to large simulations because the results it produces are too conservative, i.e. it predicts the greatest possible uncertainty in the output. Also, its ease of implementation is out-weighed by the lack of information it provides compared with probabilistic methods.

6.1.2 Uncertainty Propagation Using Sensitivity Derivatives

This deterministic technique has been in use for many years. Its aim is to quantify the sensitivity of the output to a variation in one of the inputs. In this way the total uncertainty in the output is apportioned between the various sources of uncertainty. It identifies which sources of uncertainty have a significant effect on the output and which sources can be ignored. As such, it is often a useful preliminary step performed before a computationally expensive probabilistic uncertainty analysis.

6.1.3 Possibilistic Methods - Fuzzy Logic

Fuzzy logic can be used to estimate the uncertainty in the output when the input parameter uncertainties are characterized by membership functions. Fuzzy logic calculates approximate behaviour of the system using models based on inexact, incomplete, or unreliable data. The membership function represents the degree of membership of the fuzzy variable within a fuzzy set. Uncertainty analysis using membership functions is sometimes referred to as "possibilistic", to contrast this approach with the probabilistic approach using probability density functions to characterise the uncertainty in the input parameters. The outputs are fuzzy sets, which are calculated using fuzzy algorithms based on a collection of rules, in the form of conditional statements. A range of techniques may be used to de-fuzzify the fuzzy output set to give a single (crisp) value. Although fuzzy logic is appealing because of its simplicity, many aerospace applications require a high level of accuracy which makes it necessary to employ probabilistic methods wherever it is possible to do so.

6.2 Probabilistic Methods

Probabilistic methods can be divided into "statistical" and "non-statistical" methods.

- Statistical methods use a large number of values of the input variables to calculate values repeatedly for the output variables. A sampling method is used to generate the input values from a distribution. Statistics, such as the mean and variance, of the output values can then be calculated. The classic example of a statistical method is the Monte Carlo method. These methods are computationally expensive.
- Non-statistical methods use an analytical treatment of the uncertainty and are much less computationally expensive. Two examples are the "perturbation" or "moment" methods and "stochastic differential equation" (SDE) methods.

6.2.1 Monte Carlo Methods

6.2.1.1 Basic Monte Carlo Method

Monte Carlo methods, which are also called "statistical simulation methods", can be loosely defined to include any method that utilizes sequences of random numbers to perform the simulation. Although they have been used for hundreds of years, and notably by Lord Kelvin in a paper published in 1901 (Kelvin 1901), the name "Monte Carlo" was given to these methods in 1944 by the group of elite mathematicians and physicists who were responsible for developing both the first atomic bomb and the first electronic digital computers. The analogy of Monte Carlo methods to games of chance is a good one, where the "game" is a physical system, the "winner" is the scientist and his "prize" is not money but rather a solution to some problem.

The procedure for the basic (or crude) Monte Carlo method involves three steps:

1. for each input variable, generate a set of values by randomly sampling the known or assumed probability density function
2. for each set of random input data execute a deterministic mathematical model and aggregate the output data
3. use the statistics of the output data set (mean, variance, skewness, kurtosis, etc.) to define its probability density function

It is important to note that while the Monte Carlo method converges to the exact stochastic solution as the number of samples goes to infinity, the convergence of the mean error estimate is slow because the standard deviation of the mean scales inversely with the square root of the number of samples. Hence thousands or millions of data samples may be required to get the required accuracy, and for each sample the deterministic mathematical model used in step 2 will need to be executed.

The Monte Carlo method has been applied to some very small CFD simulations in order to demonstrate the method (Walters and Huyse 2002). The meshes they used

were between 1,000 and 5,000 cells. They simulated oblique shocks and expansions with uncertainty in the Mach number and flow turning angle; supersonic flow over a thin airfoil with uncertain thickness; and incompressible flow over a flat plate with uncertainty in the viscosity.

The basic Monte Carlo method is very computationally expensive. To reduce the computational expense, modifications of the basic Monte Carlo method have been developed. These efficiency improvements are known as variance reduction techniques. Two of the modified Monte Carlo methods are called "Importance Sampling" and "Latin Hypercube Sampling". They are popular because they are easy to implement and reduce the computation time while still providing the required accuracy in the particular situations described below.

6.2.1.2 Importance Sampling Monte Carlo Method

Importance Sampling is one of the variance-reducing modifications to the Monte Carlo method. The fundamental idea is that the sampling process is deliberately distorted or biased. Instead of taking samples at random from a PDF, a "sampling PDF" is designed to take more samples from a region of importance. Importance sampling is particularly valuable because the largest gains are possible in some of the most difficult simulation applications, those involving the simulation of rare events such as system failures. This is very important in risk and reliability analysis. A sampling distribution designed to take more samples in the region where events are rare reduces the number of simulations required to observe an adequate number of such events. In order to "unbias" the results, to counteract the bias introduced by using the sampling distribution in place of the true distribution, a weighting function is applied to each observation. The weighting function is inversely proportional to the relative likelihood of observing an event under the sampling distribution compared to the likelihood of observing the event under the true distribution. Thus, if a given event is twice as likely to be observed, it is given half the weight when it is observed.

Reductions in computation time of many orders of magnitude are possible by using Importance Sampling instead of the basic Monte Carlo method with the same accuracy when rare events are involved. A simulation which takes 10,000 minutes using the basic Monte Carlo method may take only one minute using Importance Sampling. Unfortunately, in applications which do not involve rare events the design of the sampling distribution can be very difficult with the result that the Importance Sampling method may increase the simulation time or fail.

6.2.1.3 Latin Hypercube Monte Carlo Method

In the Latin Hypercube sampling method the selection of sample points is highly constrained. This enables the statistics of the sample (mean, variance, etc.) to resemble the statistics of the probability density function (PDF) being sampled with the same accuracy for a smaller sample size than is required for the basic Monte Carlo sampling

method. For a single uncertain input parameter, instead of taking N samples at random from the complete PDF, the range of probable values is partitioned into N segments of equal probability. That is, each segment corresponds to an equal area under the PDF curve. An example is shown in Figure 2 for a single parameter range divided into N = 10 segments. Then one random sample is taken from within each segment, yielding N samples.

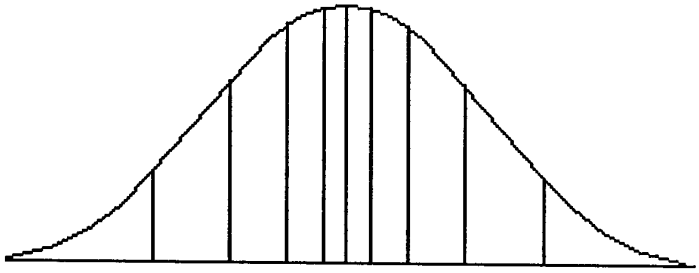


Figure 2: a single parameter range divided into 10 segments of equal probability

For M parameters and N segments, the whole parameter space is partitioned into N^M cells, each having equal probability. For 2 parameters the space is a square, for 3 parameters it is a cube, and for more than 3 parameters it is called a "hypercube", which is where the second half of the "Latin Hypercube" name comes from. For example, for the case of M = 3 parameters divided into N = 4 segments each, the parameter space is divided into 4 x 4 x 4 = 64 cells. The cells can be identified by an M dimensional vector of indices where each index can take values from 1 to N. When M = 3 and N = 4 the cell label will look like (3,1,4). It is then necessary to select N cells, from each of which one random sample will be taken. The N cells must be selected in such a way that each segment from each parameter range is selected once and only once.

This corresponds to the idea of Latin Squares which explains where the first half of the "Latin Hypercube" name comes from. Latin Squares have been used to design efficient statistical experiments since at least 1925. In a Latin Square the numbers 1 to n each occur n times in an nxn matrix in such a way that in any row or column no number is repeated. An example for n = 4 is shown in Figure 3.

2	1	4	3
1	2	3	4
4	3	2	1
3	4	1	2

Figure 3: A Latin Square for n = 4

An example of the use of a Latin Square to design an efficient experiment would be if we wanted to test four different CFD codes to determine whether there was any

performance difference between them. Suppose that we also know that our testing environment includes two sources of variability. For example, the performance data we can measure for each CFD code may depend not just on the code itself but also on the behaviour of the individual analyst running the code, and the type of physical system being simulated. The most efficient way to arrange the series of 16 test runs is according to a Latin Square. If we design the experiment using a Latin Square it would be of order four, that is four rows and four columns, giving a total of 16 different tests. One of the many possible arrangements is as shown in Figure 3. The numbers from 1 to 4 correspond to the four CFD codes. Each row could be assigned to a different type of physical system being simulated. Each column could be assigned to a different analyst. The Latin Square ensures that each row and column contains each CFD code once and only once. This means that each code would be used once by each analyst and once with each type of simulation. In order to test the hypothesis that there is a difference between the four CFD codes it is not necessary to perform tests with the 64 possible combinations of every code with every analyst with every type of simulation. Statistical "analysis of variance" methods can then be used with the data from the 16 tests to determine how much of the variation in the results could be attributed to the analysts, and how much to the type of simulation; and then to determine whether there is any significant performance variation among the four CFD codes.

In the Latin Hypercube method with M parameters and N segments the easiest way to select N cells from the N^M cells is by randomly selecting one segment from each parameter PDF, and repeating this step N times using selection without replacement. This produces the required N cell labels with each index taking each of its possible values once. If each cell label is thought of as a vector with M components identifying one segment for each parameter, then the N vectors can be assembled into an NxM matrix. This purely random method of selecting the combinations of segments that will define the cells from which to take samples will, in general, not result in zero correlation between the values in the columns of the matrix. In other words, large values of one parameter may tend to be paired with large values of another parameter. An optional further step in the Latin Hypercube sampling method is deliberately to choose the cells from which samples will be taken in such a way that the correlation between the values in the columns of the matrix is as small as possible.

The advantage of the Latin Hypercube sampling method is that the random samples are generated from all the ranges of possible values. While this method ensures complete coverage of the full range of the input parameters it does not give accurate information about the tails of the output probability density functions. It requires about five times fewer samples than the basic Monte Carlo method, and is extremely useful when the response of the system to frequent small disturbances is of more interest than the response to the very unlikely extreme events.

The Latin Hypercube Sampling method is classified as a "constrained sampling" or "stratified sampling" method. (Iman RL, Helton JC, and Campbell JE, 1981). It was developed by McKay, Conover, and Beckman (1979).

6.2.1.4 *Response Surface Monte Carlo Method*

Pratt & Whitney developed a probabilistic design system (Fox 1994) incorporating two different methods depending on whether or not the computer codes being used were computationally expensive to run. For the codes that could be run quickly they used the Monte Carlo method of randomly varying the inputs and calculating the mean and variance of the calculated outputs. For slower codes they used a second order response surface model with Box-Behnken design experiments, followed by a Monte Carlo simulation. The Box-Behnken design experiments use just three values of each uncertain input variable: low, nominal and high. With six variables the Box-Behnken design means the computer code has to be run only 49 times. A second order response surface regression equation is then fitted to the output variables from the 49 runs. A Monte Carlo simulation is simply performed by selecting a random sample value of each input variable according to its PDF, and then calculating the corresponding output values from the response surface equation. Several thousand Monte Carlo samples can be evaluated in a very short time.

6.2.1.5 *Random Number Generators for Monte Carlo Simulations*

Generating "good" random numbers can be a major problem in Monte Carlo simulations. Many basic random number generators (RNGs) supplied with computers are not sufficiently random. Even some RNGs with adequate randomness properties are not good enough for Monte Carlo simulation because their period of repetition is too small. The period must be much greater than the number of random numbers used in the simulation, or else the results can be incorrect.

Multiplicative linear congruential generators (MLCG) using 32-bit integers have a period of at most $2^{31} \cong 10^{10}$. This many random numbers can be generated in a matter of seconds on a modern workstation. On the other hand, using 64-bit words or combining two 32-bit MLCGs can give periods $\cong 10^{18}$. These are adequate for large simulations requiring a Gigaflop-year of computer time. Other generators (lagged Fibonacci, etc.) have even longer periods. These or combined 64-bit MLCGs will be required for Teraflop-year simulations. A Pentium-4 machine is capable of calculating at a rate of roughly 2 Gigaflops. At this rate a Gigaflop-year size simulation would take six months to run. This explains why parallel computers with hundreds of nodes are not uncommon.

6.2.2 Moment Methods

The moments of a probability distribution are its defining features.

- the first moment is the mean
- the second moment is the variance
- the third moment is the skewness
- the fourth moment is the kurtosis (peakiness)

A normal or Gaussian probability distribution is completely defined by its first two moments, its mean and variance. Other types of probability distribution are not so simple. Moment methods ignore the higher order moments and approximate the probability density functions which characterise the uncertainty in the input parameters by their first two moments.

Moment methods are very simple, convenient and widely used approximate methods for estimating the moments of the output from the moments of the uncertain input parameters. The moments of the output are calculated from truncated Taylor series expansions about the mean value of the input. The four basic versions of this method are explained below for the case of one single uncertain input parameter. Extension of the methods to multiple input parameters is straightforward using Taylor series expansions involving multiple variables. These methods are also called "perturbation methods".

6.2.2.1 First-Order, First Moment (FOFM)

The first-order accurate Taylor series expansion to estimate the mean (first moment) of the output, $E_{\text{first order}}[y(x)]$, is simply the output ($y$) calculated using the first moment of the input $(\bar{x})$. In other words, an approximation to the mean value of the output from a CFD code can be obtained by running the code using the mean value of the uncertain input parameter.

$$E_{\text{first order}}[y(x)] = y(\bar{x})$$

This is called the "deterministic solution". It is important to remember the counter intuitive fact that this only gives an approximate value of the mean of the output and not the exact value. It assumes that the probability distribution of the input parameter is symmetric about the mean, which may not be true.

6.2.2.2 Second-Order, First Moment (SOFM)

The second-order accurate Taylor series expansion to estimate the mean (first moment) of the output improves the estimate by adding a second term. This term comprises half the variance of the input multiplied by the second partial derivative of the output with respect to the input (the second sensitivity derivative) evaluated at the mean value of the input. For some problems this higher order correction term can be significant.

$$E_{\text{second order}}[y(x)] = y(\bar{x}) + \frac{1}{2}\text{Var}(x)\left.\frac{\partial^2 y}{\partial x^2}\right|_{\bar{x}}$$

6.2.2.3 *First-Order, Second Moment (FOSM)*

The first-order accurate Taylor series expansion to estimate the variance (second moment) of the output is equal to the variance of the input parameter multiplied by the square of the first sensitivity derivative evaluated at the mean value of the input.

$$\mathrm{Var}_{\text{first order}}[y(x)] = \mathrm{Var}(x)\left(\left.\frac{\partial y}{\partial x}\right|_{\bar{x}}\right)^2$$

6.2.2.4 *Second-Order, Second Moment (SOSM)*

The second-order accurate Taylor series expansion to estimate the variance (second moment) of the output has a higher order correction term involving the square of the second sensitivity derivative evaluated at the mean value of the input.

$$\mathrm{Var}_{\text{second order}}[y(x)] = \mathrm{Var}(x)\left(\left.\frac{\partial y}{\partial x}\right|_{\bar{x}}\right)^2 + \frac{1}{2}\left(\mathrm{Var}(x)\left.\frac{\partial^2 y}{\partial x^2}\right|_{\bar{x}}\right)^2$$

6.2.2.5 *Summary of Moment Methods*

Moment methods are popular because of their low computational expense. They will not give good results unless the probability distribution of the uncertainty in the input parameters is close to Gaussian because they ignore the higher order moments. It is essential to have an efficient procedure for calculating the sensitivity derivatives. Huyse has applied them to CFD simulations at NASA Langley Research Centre for airfoil shape optimisation under uncertainty (Huyse 2001).

6.2.3 Stochastic Differential Equation Methods

Stochastic Differential Equation (SDE) methods have been used to model the uncertainty in CFD simulations by adding stochastic (random) variables to the deterministic CFD equations (Mathelin, L., et al.2003). These methods are based on the work of Weiner in 1938 (Weiner 1938) and have been used for structural mechanics problems. They have only been applied to CFD problems quite recently. They provide the information about the higher moments of the results at less computational expense than Monte Carlo methods.

6.2.3.1 *Polynomial Chaos (PC)*

Polynomial Chaos uses a spectral representation of the uncertainty which is then decomposed into separate deterministic and random components. In the Polynomial Chaos method each variable in the equations of the mathematical model, such as pressure or velocity, is expanded into an infinite series using Hermite polynomials. In theory the series is infinite, but for practical problems it must be truncated to a finite

series. The first term in the series represents the mean value of the variable. The second term represents the Gaussian random fluctuations around the mean value. The third and higher terms represent non- Gaussian random fluctuations. In this way the random behaviour is decomposed into a finite number of orthogonal modes of fluctuation. This concept is similar to the modes of vibration of a mechanical system which occur at particular points in the frequency spectrum. A Galerkin projection can then be used to produce a set of deterministic equations that can be solved with conventional methods.

Xiu and Karniadakis (2003) used a Polynomial Chaos method to solve a transient heat conduction problem. They obtained the statistics of the result (mean, variance, etc.) 500 times faster with this method than with a Monte Carlo method. The Monte Carlo method required over 20,000 samples to achieve the same accuracy as the polynomial chaos method with 35 samples. They noted that the efficiency of the chaos expansion is problem specific and requires further research.

Lucor and his colleagues applied the polynomial chaos method to the incompressible Navier-Stokes equations for flow around an elastically mounted cylinder for laminar flow at a low Reynolds number (Lucor, D. et al. 2003). The damping coefficient and the spring factor of the structure were both assumed to be random variables. The free structure, excited by the vortex shedding of the flow which is initially deterministic, produces a random response. Therefore, the position of the boundary of the cylinder becomes stochastic. This random boundary affects the flow domain, and consequently the flow itself becomes a stochastic process. They found that "the regions of the flow domain with high uncertainty were the shear layers and the near-wake of the cylinder, which are of course the regions of utmost physical interest." Speed-up factors from 1000 to more than 100 000 compared to Monte Carlo simulation were observed, depending on the problem. They discuss the limitations of Hermite expansions in applications with turbulence and suggest a possible solution using Askey polynomials called "Generalised Polynomial Chaos". They conclude that more research is needed before Polynomial Chaos will be incorporated into CFD simulations.

7. Conclusions and Recommendations

CFD can produce critical data about engineering systems in situations where it is too difficult, dangerous or expensive to perform experiments. CFD is a complex process which requires expertise and experience in the analyst to achieve good results because it may include many sources of error, uncertainty and variability. Estimation of the uncertainty from all sources in a CFD simulation requires the three distinct processes of verification, validation and uncertainty analysis.

Often the uncertainty and variability in the input data for a CFD simulation can be characterised by probability density functions. In this case, the uncertainty can be

propagated through the analysis to the results using a range of probabilistic uncertainty analysis methods. The use of probabilistic techniques in CFD simulations has recently become a very active area of research.

Moment methods have been widely used because they are quick and simple when only the first and second statistical moments of the results are needed (i.e. mean and variance), but they can be used only when the uncertainty distribution can be assumed to be close to Gaussian.

The basic Monte Carlo method provides the complete information about the higher statistical moments of the results but it is very computationally expensive. In order to reduce this expense modified Monte Carlo methods such as Latin Hypercube Sampling or Importance Sampling have been developed. In spite of these improvements in sampling methods, the Monte Carlo method remains prohibitively computationally expensive for CFD simulations using meshes large enough to solve practical engineering problems.

An approximate Monte Carlo method which substitutes a response surface for the CFD model is computationally cheap and has been used in practice by an engine manufacturer. It is, at present, the most economical option for obtaining the full statistical information about the simulation results.

Stochastic Differential Equation methods provide the complete information about the higher statistical moments at less computational expense than Monte Carlo methods. Therefore, if this information is needed, and if the approximation involved in the response surface Monte Carlo method is not acceptable, the SDE methods are the best alternative. Research in the application of these methods to CFD simulation is just beginning. It will be some time before they are widely adopted.

It is well understood in the CFD community that for certain types of problems it is not possible to achieve a highly accurate solution. Nevertheless, in the hands of experts, CFD has produced a level of insight into the behaviour of physical systems which it is not feasible to obtain in any other way. There is a growing recognition of how important it is to quantify error, uncertainty and variability in computational simulation. The knowledge and practical experience of CFD experts are being drawn upon to establish procedures that quantify both the inaccuracies we know we have introduced to a simulation, and also those things which we are not certain about or which are stochastic (random) in nature. A quotation that is frequently repeated in the field of uncertainty analysis because it reminds us that all mathematical models are only an approximation to reality is: "All models are wrong, some are useful." (Box 1980). There is no doubt that an estimate of how "wrong" a model is will increase its usefulness.

8. References

1. AIAA, "Guide for the Verification and Validation of Computational Fluid Dynamics Simulations," AIAA G-077-1998, 1998

2. AGARD (NATO) Fluid Dynamics Panel, Advisory Report No. 303 , 1994, "A Selection Of Experimental Test Cases For The Validation Of CFD Codes"

3. ANSI/ASME Performance Test Codes PTC 19.1-1985 "Measurement Uncertainty - Part 1 Instruments and Apparatus"

4. Assessment of Experimental Uncertainty with Application to Wind Tunnel Testing," AIAA Standard S-071A-1999, Revision A of the Standard, AIAA, Reston, VA, 1999, pp. 11–17.

5. Box, G. E. P., 1980, "Sampling and Bayes' Inference in Scientific Modelling and Robustness," Journal Statist. Soc. A, Vol. 143, No. A, pp. 383-430.

6. Cox, E., The Fuzzy Systems Handbook: A Practitioner's Guide to Building, Using, and Maintaining Fuzzy Systems, 2nd ed., AP Professional, Division of Academic Press, San
7. Diego, CA, 1999.

8. de Cooman, G., Ruan, D., and Kerre, E. E., eds. Foundations and Applications of Possibility Theory, World Scientific Publishing Co., Singapore, 1995.

9. Fox, E.P., 1994, "Pratt & Whitney Probabilistic Design System", Proceedings of the AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, pp.1075-1085, AIAA-94-1442-CP

10. Guan, J., and Bell, D. A., Evidence Theory and Its Applications, Vol. I, Elsevier Science Publishers B.V., Amsterdam, The Netherlands, 1991.

11. Huyse, L., 2001, "Free-Form Airfoil Shape Optimization under Uncertainty Using Maximum Expected Value and Second-Order Second-Moment Strategies", ICASE Report 2001-18, NASA/CR-2001-211020

12. Iman RL, Helton JC, and Campbell JE, (1981) "An approach to sensitivity analysis of computer models, Part I. Introduction, input variable selection and preliminary variable assessment", Journal of Quality Technology 13:174-183,

13. Iman RL, Helton JC, and Campbell JE, (1981) "An approach to sensitivity analysis of computer models, Part II. Ranking of input variables, response

surface validation, distribution effect and technique synopsis", Journal of Quality Technology 13:232-240,

14. International Organisation for Standards, 1995, "Guide to the Expression of Uncertainty in Measurement"

15. Kappas, J. (2002). "Review of Risk and Reliability Methods for Aircraft Gas Turbine Engines." DSTO-TR-1306

16. Kelvin, L., 1901, "Nineteenth Century Clouds Over the Dynamical Theory of Heat and Light", Phil. Mag., 2, pp.1-40.

17. Kozine, I., "Imprecise Probabilities Relating to Prior Reliability Assessments," 1st International Symposium on Imprecise Probabilities and Their Applications, Ghent, Belgium, 1999.

18. Lincoln, J. W. (1985). "Risk Assessment of an Aging Military Aircraft." Journal of Aircraft 22(8)

19. Lucor, D., D. Xiu, et al. (2003), "Predictability and uncertainty in CFD." Int. J. Numer. Meth. Fluids 43(5): 483-505.

20. MARNET-CFD (2000)
https://pronet.wsatkins.co.uk/marnet/guidelines/guide.html

21. Mathelin, L., M. Y. Hussaini, et al. (2003), "Uncertainty Propagation for Turbulent, Compressible Flow in a Quasi-1D Nozzle Using Stochastic Methods." 16th AIAA Computational Fluid Dynamics Conference, Orlando, Florida, AIAA 2003-4240

22. McKay, M.D., Conover, W.J., and Beckman, R.J. (1979). "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," Technometrics, 221, 239-245.

23. Mehta, U.B., "Guide to Credible Computer Simulations of Fluid Flows," AIAA Journal of Propulsion and Power, Vol. 12, No. 5, September-October 1996, pp. 940-948. (Also AIAA Paper 95-2225).

24. Nagpal, V. K., R. Rubinstein, et al. (1989) "Probabilistic structural analysis to quantify uncertainties associated with turbopump blades", AIAA Journal 27(6): 809-813

25. Oberkampf, W. L., et al. (2000). "Estimation of Total Uncertainty in Modeling and Simulation." Sandia Report SAND2000-0824.

26. Roache, P.J., Verification and Validation in Computational Science and Engineering, Hermosa Publishers, Albuquerque, New Mexico, 1998.

27. Richardson, L.F. (1927), "The Deferred Approach to the Limit", Transactions of the Royal Society of London, Series A, Vol. 226, pp. 299 - 361.

28. Tong, Y. C. (2001). "Literature Review on aircraft structural risk and reliability analysis." DSTO-TR-1110.

29. Veach, Eric, "Robust Monte Carlo Methods for Light Transport Simulation", Ph.D. dissertation, Stanford University, December 1997

30. Walters, R.W., and Huyse, L., 2002, "Uncertainty Analysis for Fluid Mechanics with Applications", NASA/CR-2002-211449, ICASE Report No. 2002-1

31. Weiner, N. 1938, "The Homogenous Chaos", American Journal of Mathematics, Vol. 60, pp.897-936

32. Xiu, D. and G. E. Karniadakis (2003), "A new stochastic approach to transient heat conduction modeling with uncertainty." International Journal of Heat & Mass Transfer 46(24): 4681-4693

33. Yang, J. N. and S. Chen (1985). "Fatigue reliability of gas turbine engine components under scheduled inspection maintenance." Journal of Aircraft 22(5)

34. Zang, T. A., et al. (2002). "Needs and Opportunities for Uncertainty-Based Multidisciplinary Design Methods for Aerospace Vehicles." NASA/TM-2002-211462

DISTRIBUTION LIST

Probabilistic Methods for the Quantification of Uncertainty and Error in Computational Fluid Dynamics Simulations

John Faragher

AUSTRALIA

DEFENCE ORGANISATION

	No. of copies
Task Sponsor	
DGTA-ESI1A	1
S&T Program	
Chief Defence Scientist	shared copy
FAS Science Policy	
AS Science Corporate Management	
Director General Science Policy Development	
Counsellor Defence Science, London	Doc Data Sheet
Counsellor Defence Science, Washington	Doc Data Sheet
Scientific Adviser to MRDC, Thailand	Doc Data Sheet
Scientific Adviser Joint	1
Navy Scientific Adviser	1
Scientific Adviser - Army	1
Air Force Scientific Adviser	1
Scientific Adviser to the DMO M&A	1
Scientific Adviser to the DMO ELL	1
Platforms Sciences Laboratory	
Chief of AVD	Doc Data Sht & Dist List
Research Leader (Sam Fisher)	pdf
Task Manager (Ross Antoniou)	pdf
Bryon Wicks	pdf
Author: John Faragher	pdf
Chun Wang	pdf
Jianfu Hou	pdf
Wyman Zhuang	pdf
Joanna Kappas	pdf
Yu Chee Tong	pdf
DSTO Library and Archives	
Library Fishermans Bend	Doc Data Sheet
Library Edinburgh	1
Defence Archives	1

Capability Development Group

Director General Maritime Development	Doc Data Sheet
Director General Capability and Plans	Doc Data Sheet
Assistant Secretary Investment Analysis	Doc Data Sheet
Director Capability Plans and Programming	Doc Data Sheet
Director Trials	Doc Data Sheet

Chief Information Officer Group

Deputy CIO	Doc Data Sheet
Director General Information Policy and Plans	Doc Data Sheet
AS Information Strategy and Futures	Doc Data Sheet
AS Information Architecture and Management	Doc Data Sheet
Director General Australian Defence Simulation Office	Doc Data Sheet
Director General Information Services	Doc Data Sheet

Strategy Group

Director General Military Strategy	Doc Data Sheet
Director General Preparedness	Doc Data Sheet
Assistant Secretary Strategic Policy	Doc Data Sheet
Assistant Secretary Governance and Counter-Proliferation	Doc Data Sheet

Navy

SO (SCIENCE), COMAUSNAVSURFGRP, NSW Doc Data Sht & Dist List

Maritime Operational Analysis Centre, Building 89/90 Garden Island Sydney NSW Doc Data Sht & Dist List

Deputy Director (Operations) & Deputy Director (Analysis) Doc Data Sht & Dist List

Director General Navy Capability, Performance and Plans, Navy Headquarters Doc Data Sheet

Director General Navy Strategic Policy and Futures, Navy Headquarters Doc Data Sheet

Air Force

SO (Science) - Headquarters Air Combat Group, RAAF Base, Williamtown NSW 2314 Doc Data Sht & Exec Summ

Army

ABCA National Standardisation Officer, Land Warfare Development Sector, Puckapunyal e-mailed Doc Data Sheet

SO (Science) - Land Headquarters (LHQ), Victoria Barracks NSW Doc Data & Exec Summ

SO (Science), Deployable Joint Force Headquarters (DJFHQ) (L), Enoggera QLD Doc Data Sheet

Joint Operations Command

Director General Joint Operations	Doc Data Sheet
Chief of Staff Headquarters Joint Operations Command	Doc Data Sheet
Commandant ADF Warfare Centre	Doc Data Sheet
Director General Strategic Logistics	Doc Data Sheet

Intelligence and Security Group

DGSTA Defence Intelligence Organisation	1
Manager, Information Centre, Defence Intelligence Organisation	1 (PDF)
Assistant Secretary Capability Provisioning	Doc Data Sheet
Assistant Secretary Capability and Systems	Doc Data Sheet
Assistant Secretary Corporate, Defence Imagery and Geospatial Organisation	Doc Data Sheet

Defence Materiel Organisation

Deputy CEO	Doc Data Sheet
Head Aerospace Systems Division	Doc Data Sheet
Head Maritime Systems Division	Doc Data Sheet
Chief Joint Logistics Command	Doc Data Sheet

Defence Libraries

Library Manager, DLS-Canberra	Doc Data Sheet
Library Manager, DLS - Sydney West	Doc Data Sheet

OTHER ORGANISATIONS

National Library of Australia	1
NASA (Canberra)	1

UNIVERSITIES AND COLLEGES

Australian Defence Force Academy	
Library	1
Head of Aerospace and Mechanical Engineering	1
Serials Section (M list), Deakin University Library, Geelong, VIC	1
Hargrave Library, Monash University	Doc Data Sheet
Librarian, Flinders University	1

OUTSIDE AUSTRALIA

INTERNATIONAL DEFENCE INFORMATION CENTRES

US Defense Technical Information Center	2
UK Defence Research Information Centre	2
Canada Defence Scientific Information Service	e-mail link to pdf
NZ Defence Information Centre	1

INFORMATION EXCHANGE AGREEMENT PARTNERS

National Aerospace Laboratory, Japan	1
National Aerospace Laboratory, Netherlands	1

SPARES	5
Total number of copies:	**29 printed and 10 PDF**

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION
DOCUMENT CONTROL DATA

1. PRIVACY MARKING/CAVEAT (OF DOCUMENT)

2. TITLE

Probabilistic Methods for the Quantification of Uncertainty and Error in Computational Fluid Dynamics Simulations

3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION)

Document	(U)
Title	(U)
Abstract	(U)

4. AUTHOR(S)

John Faragher

5. CORPORATE AUTHOR

Platforms Sciences Laboratory
506 Lorimer St
Fishermans Bend Victoria 3207 Australia

6a. DSTO NUMBER	6b. AR NUMBER	6c. TYPE OF REPORT	7. DOCUMENT DATE
DSTO-TR-1633	AR-013-229	Technical Report	October 2004

8. FILE NUMBER	9. TASK NUMBER	10. TASK SPONSOR	11. NO. OF PAGES	12. NO. OF REFERENCES
2004/1029533/1	AIR 03/121	DGTA	30	34

13. URL on the World Wide Web

http://www.dsto.defence.gov.au/corporate/reports/DSTO-TR-1633.pdf

14. RELEASE AUTHORITY

Chief, Air Vehicles Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved for public release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SA 5111

16. DELIBERATE ANNOUNCEMENT

No Limitations

17. CITATION IN OTHER DOCUMENTS Yes

18. DEFTEST DESCRIPTORS

Gas turbine engines, Aircraft engines, Reliability (engineering), Risk assessment, Military aircraft, Probability theory

19. ABSTRACT

The life of engine components is determined by a combination of the material properties and the applied stresses and temperatures. As a consequence of variability in these parameters, the component life is not fixed (deterministic) but stochastic (random) and may be characterised by a probability density function (PDF). In order to reduce the cost of ownership of ADF aircraft these PDFs need to be determined as accurately as possible. Probabilistic techniques offer significant potential for accurate and realistic estimates of component lives by quantifying stochastic elements of an analysis rather than introducing excessive conservatism to allow for them. This report examines the feasibility of using a probabilistic approach for modelling the component temperatures in an engine using CFD (Computational Fluid Dynamics).