

Development Strategies for Uncertainty Quantification to Enable Aircraft Certification by Analysis

John A. Schaefer*

Boeing Research & Technology, Saint Louis, Missouri, 63166

Philipp Bekemeyer†

*German Aerospace Center (DLR), Institute of Aerodynamics and Flow Technology,
38108 Braunschweig, Germany*

Casey L. Denham‡

NASA Langley Research Center, Hampton, Virginia, 23681

Simon Coggon§

Airbus Operations Ltd Filton, Bristol, BS347PA, United Kingdom

Aircraft Certification by Analysis (CbA) is a broad term that describes the process by which engineering analysis tools are used to supplement flight testing to demonstrate compliance with regulatory requirements. For each requirement that is to be met through this process, the analysis tools must be sufficiently accurate to ensure that an equivalent level of safety is obtained as if the aircraft to be certified had been flight tested for the requirement. In recent decades, the ability of analysis tools to accurately predict aircraft performance has improved dramatically, leading to increasing interest and limited early adoption of Certification by Analysis across the aerospace industry. However, rigorously demonstrating the credibility of analysis tools remains a challenge for many types of certification requirements, yet is necessary for widespread adoption of CbA approaches in the future. The field of uncertainty quantification is well-suited to address this problem, but will require substantial advances in order to achieve the ambitious industry goals for Certification by Analysis. The objective of this position paper is to highlight gaps in the current technical and programmatic state of uncertainty quantification. Four main categories – Mindset & Awareness, Tools & Capabilities, Data and Benchmarks for V&V, and Applied Research & Established Processes – will be introduced. Each of these categories contains several shortcomings and impediments which, if overcome, will enable more widespread possibilities for Certification by Analysis in the 2040-2050 timeframe.

I. Introduction

As of today, aircraft certification, no matter if for minor modifications or completely new parts or designs, is fundamentally built upon flight testing. The ambition of Certification by Analysis (CbA) is to selectively reduce the amount of certification flight testing required for new aircraft by supplementing it with analysis-based results, while maintaining an equivalent level of safety as if a flight test had been performed.¹ This goal, if achieved, will enable faster time-to-market, reduced development costs, and more innovative designs, which will in turn lead to higher-quality aircraft for the airlines and a better overall experience for the flying public. CbA is fundamentally built upon flight modelling, which is a term that covers a variety of

*Aerodynamics Engineer, The Boeing Company, Senior Member AIAA

†Team Leader, C²A²S²E Department, Member AIAA

‡Aerospace Technologist - Aerospace Vehicle Design and Mission Analysis, Aeronautics Systems Analysis Branch, Member AIAA

§Expert in Uncertainty Quantification and Management, Airbus

methods, including analyses based on wind tunnel results, numerical methods such as Computational Fluid Dynamics (CFD), Computational Structural Mechanics (CSM), and Computational Aeroelasticity (CA), as well as Flight Dynamics Simulation (FDS).

The idea of reducing physical testing through the usage of analysis methods is neither new nor limited to aircraft certification. In fact, various activities are ongoing in different industries with the same or similar ambitions, even if the naming convention differs.^{2,3} One example is the nuclear industry, where certification is already commonly done by analysis, due to the high risks and costs associated with testing.² Simulation in general is not new to the aerospace industry, and in fact contributes significantly to every modern aircraft program throughout the design cycle, using tools from many disciplines at varying levels of fidelity. Within recent years, increased performance and availability of computational resources combined with the development of sophisticated simulation tools has increased the desire and plausibility of using simulations more rigorously during aircraft certification.

In 2021 the AIAA published the *Recommended Practice: When Flight Modelling Is Used to Reduce Flight Testing Supporting Aircraft Certification*, in which a set of practices are outlined when flight modelling is being developed, proposed, and used to reduce flight testing relative to established aircraft certification practices.¹ In total six tasks are recommended, including:

1. Configuration and process management
2. Verification of the models
3. Verification that the models were applied correctly
4. Validation of the models
5. Justification of analysis adequacy in recognition of potential modelling error and/or uncertainty
6. A summary assessment of applicability for showing compliance

Most tasks are closely linked to established development and application processes for models (e.g., verification), and various activities are ongoing at different institutions and at different levels worldwide to support these efforts. However, the field of uncertainty quantification (UQ) mentioned in the fifth item above is less established, even though UQ plays a fundamental role for successfully addressing all six CbA tasks. It is fair to say that significant advances in the field of UQ and its integration with aerospace models and processes are necessary to achieve the ambitious goals of CbA. This has already been acknowledged within the Recommended Practice¹ as well as the NASA Guide on CbA,⁴ in which an explicit UQ technology development track is present.

In the next section, a short overview of UQ is provided, including a few common terms and concepts in the field. References are provided for the interested reader to follow up for a more in-depth introduction to UQ. Section III outlines UQ technologies and programmatic impediments which will need to be addressed in order to pave the way for CbA. These challenges are clustered into four main categories – Mindset & Awareness, Tools & Capabilities, Data & Benchmarks for V&V, and Applied Research & Established Processes – and detailed shortcomings for each category are presented. Finally, general conclusions are given in Section IV to highlight the most important points going forward for establishing UQ for CbA.

II. Overview of UQ

Uncertainty Quantification is the process of characterizing all major sources of uncertainty in the model and experiment, and quantifying their effect on the analysis outcomes. UQ is a closely related activity to verification and validation (V&V) and is essential for verifying and validating analysis models. For practical engineering purposes, the goal of UQ, along with V&V, is to enable modelers and analysts to make justifiable statements about the accuracy/uncertainty and associated degree of credibility in their analysis-based predictions. In this context, UQ can be defined⁵⁻⁷ as the

- Identification (Where are the uncertainties?),
- Characterization (What form are they, and what are their mathematical descriptions?),
- Propagation and Aggregation (How do they combine to determine total uncertainty in the analysis results?), and
- Analysis (What are their impacts and implications?)

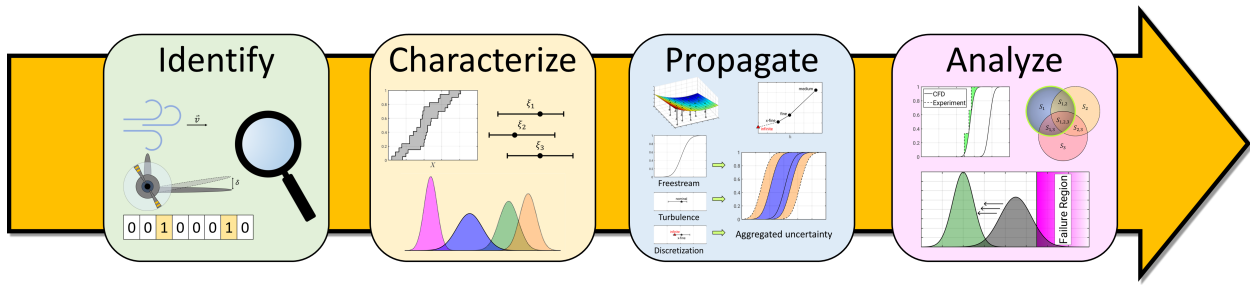


Figure 1. The four steps of uncertainty quantification.

of uncertainties in analysis models. These steps are shown graphically in Fig. 1. The outcome of UQ is the enablement of modelers and analysts to make more informed statements about the uncertainty and associated degree of credibility they have in their analysis-based predictions, as compared to the practice of performing deterministic analysis (i.e., not considering uncertainties).

There are many important considerations to be made during each of the four steps of UQ, and a companion paper to the aforementioned AIAA Recommended Practice discusses these ideas in the context of CbA.⁸ One key consideration is that, in general, there are two forms of uncertainty – aleatory and epistemic – which can be defined as:

- Aleatory: A type of uncertainty which is due to inherent, irreducible chance. Aleatory uncertainty is probabilistic in nature, and most commonly described by a probability density function (PDF) or cumulative distribution function (CDF).
- Epistemic: A type of uncertainty which is due to lack of knowledge and is potentially reducible. This type of uncertainty could be probabilistic (with an unknown PDF/CDF) or non-probabilistic (e.g., an interval with lower and upper bounds, but no probabilistic information assumed between).

Classical statistics and Bayesian methods deal entirely with aleatory uncertainty; however many problems in CbA include epistemic uncertainty as well. It should be noted that a uniform distribution is not equivalent to an epistemic interval because the former assumes that probability is evenly distributed between bounds, whereas the latter makes no such assumptions. This presents a major challenge for the continued development of UQ technology. An in-depth review of existing capabilities within the field of UQ in the context of CbA is beyond the scope of this paper, but the interested reader is encouraged to review other works related to this topic.^{1, 4, 5, 8}

As a point of clarification, the terms “error” and “uncertainty” are not equivalent, despite their frequent interchangeability in common vernacular. In this document, error is used according to its basic dictionary definition, consistent with the lexicon employed by Oberkampf and Roy:⁹

Error in a quantity: Deviation from the true value of the quantity.

With this definition, the error in a quantity is a deterministic value. However, if the true value of the quantity is unknown or can only be estimated, then the existence of imprecisely-known error leads to uncertainty (usually epistemic) in the quantity.

The four-step UQ framework presented herein is equally applicable to both traditional approaches, such as worst-case analysis with empirically-imposed safety factors, as it is to more rigorous and complex statistical methods. NASA STD-7009 discusses a criticality assessment strategy for choosing the level of rigor required by modeling and simulation activities, which is based on human safety and mission success criteria.⁶ A similar strategy should be adopted when choosing the level of statistical rigor required for CbA UQ activities. In fact, the rotorcraft community has also advocated for varying levels of UQ rigor in a set of preliminary guidelines for the rotorcraft certification by simulation process.¹⁰

For those in the aerospace industry who are not already familiar and comfortable with UQ concepts, finding the value of non-deterministic analysis may be difficult. To help, consider the familiar adage from George Box that “all models are wrong but some are useful.” UQ, in part, allows for an estimation of the “usefulness” of a model and can bound its “wrongness.” While a single, deterministic model result

may be simple to produce, it can provide little information as to what the true physical answer is in the presence of uncertainty. A UQ analysis, however, provides context to the model results in terms of a likelihood, distribution, or interval bounding the true result. Although a non-deterministic result may seem more generalized and less accurate, it provides information about the precision (“spread”) and credibility of models, which is essential as the industry moves towards simulation-based CbA.

Especially for complex and highly-involved models such as CFD, CSM, and CA, deterministic analysis strategies are typically employed in current practice. However, these models, no matter how complicated and advanced, do not provide the “truth” but rather an approximation of it. Improvements to complex modeling techniques are being continuously performed by distinct research groups around the world. Nevertheless, it is commonly agreed that complex analysis tools will never be perfect and errors/uncertainties have to be dealt with. Some examples of uncertainty sources for analysis models are included in Table 1.

Imposing a rigorous uncertainty strategy will require a significant alteration to current and established best practices for single-discipline simulations and will pose an even larger challenge when moving towards multi-disciplinary analysis. These issues, and many others for successful use of UQ for CbA, will be discussed in the following section.

Table 1. Sources of uncertainty in simulations relevant to CbA.

Source of uncertainty	Examples
Choice of model representation and assumption; uncertainty of simulated equations versus physical reality	Assuming potential flow versus RANS or scale-resolving CFD; using linearized versus higher-order representation of material properties
Numerical error	Discrete representation of smooth differential equations in code; finite temporal and spatial resolution; iterative convergence error of numerical schemes; finite precision arithmetic
Unknown applicability of existing engineering practices to new technologies	Laminar flow wing design; ultra-high bypass ratio versus open rotor engine design; design for novel materials; high aspect ratio, strut-based wings and blended wing bodies
Discretization of continuous operational environments to finite numbers of simulation combinations	Limited number of mass cases; discrete population of aero database points used for FDS
Uncertainties in any machine learning / AI models which have been used as part of the analysis process	Future examples may include feature detection or engine health monitoring

III. UQ Technology and Programmatic Impediments

Several technological shortcomings and programmatic impediments must be overcome to enable CbA with a special focus on UQ. In this section, these challenges will be introduced and a target audience for addressing each will be suggested. The authors hope that this provides some guidance for future UQ activities towards CbA. In total four different categories of shortcomings and impediments have been identified that condense several aspects:

- UQ Mindset & Awareness
- Tools & Capabilities
- Data & Benchmarks for V&V
- Applied Research & Established Processes

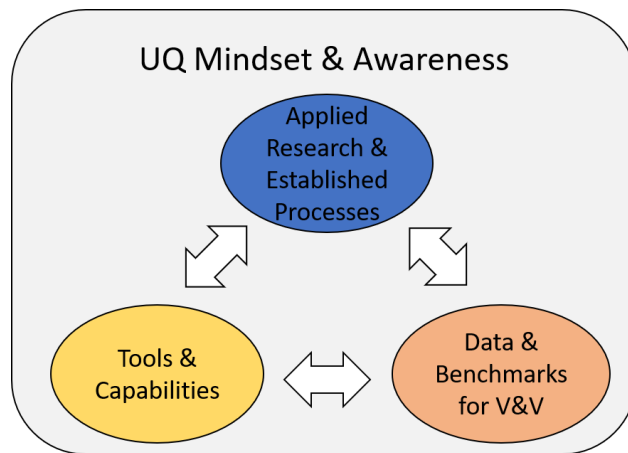


Figure 2. The four different categories for advancing UQ technology and overcoming programmatic impediments.

Fig. 2 presents these categories and outlines their connection. While the *Mindset & Awareness* category provides an overall framework of all activities, specific *Tools & Capabilities* together with *Data & Benchmarks for V&V* form the basis to enable *Applied Research & Established Processes*. A multi-directional relationship is present between the three inner categories, and clearly sorting certain aspects of UQ impediments is sometimes difficult. Moreover, the authors are aware that many activities in the field of UQ are ongoing, and note that inclusion of specific issues in the sections that follow does not mean that those issues are not being addressed at all, but rather that an additional focus on CbA might be necessary.

III.A. Mindset & Awareness

One of the major challenges for successful adoption of UQ for CbA will be to change the mindset and increase awareness of UQ technology in the aerospace industry. Perhaps the best place to start is at the collegiate level. The authors of this document surveyed the curricula of the top twelve undergraduate aerospace engineering programs in the United States (ranked according to a recent U.S. News article¹¹), and found that nearly every program required multiple courses in calculus and differential equations, but only three out of twelve programs required a course in statistics or probability theory outside of elective course offerings. Similar findings can be reported for European universities. There are many indirect consequences for CbA if undergraduates are not exposed to this material; for example: (1) students who continue on to graduate school with no background in statistics are less likely to pursue aerospace-relevant UQ research which could advance the goals of CbA; (2) students who enter the engineering workforce with no background in statistics will not be equipped to provide a rigorous measure of confidence in simulation-based results, which will be required for CbA; and (3) students with no background in statistics who eventually become engineering managers and decision makers for OEMs or regulators will not be prepared to effectively communicate with each other about alternate means of compliance based on statistical analysis. Advocates of CbA should work with universities to address this potential knowledge gap.

Another impediment to be overcome is the inclusion of unclear statistical jargon in regulation requirements. For example, consider the text below from FAA 14 CFR Part 25.933, which concerns turbojet reversing systems. Two statistically unclear terms have been italicized for demonstration:

§ 25.933(2) Each system intended for inflight use must be designed so that no unsafe condition will result during normal operation of the system, or from any failure (or *reasonably likely combination of failures*) of the reversing system, under any anticipated condition of operation of the airplane including ground operation. Failure of structural elements need not be considered if the *probability of this kind of failure is extremely remote*.

Based on this and many similar examples, OEMs must decide what is “reasonably likely” or what sort of probability constitutes “extremely remote.” FAA AC 25.1309-1A provides both qualitative and quantitative definitions of the term “extremely improbable” but does not define “extremely remote.” EASA AMC 25.1309 defines “extremely remote” as being on the order of 10^{-9} per flight hour or less, but the FAA only references

this value in the draft (not-official) version AC 25.1309-Arsenal. Even with this term clearly defined, the lack of an explicit reference to AC 25.1309 from CFR 25.933 presents a barrier for CbA UQ because it prevents statisticians and UQ practitioners from understanding the regulation without assistance from subject matter experts in the CFR. One example of a clearer regulation more compatible with CbA UQ goals is FAA 14 CFR Part 25.613, which concerns material strength properties and reads in part:

§ 25.613(b) Material design values must be chosen to minimize the probability of structural failures due to material variability. Except as provided in paragraphs (e) and (f) of this section, compliance must be shown by selecting material design values which assure material strength with the following probability:

- (1) Where applied loads are eventually distributed through a single member within an assembly, the failure of which would result in loss of structural integrity of the component, 99 percent probability with 95 percent confidence.
- (2) For redundant structure, in which the failure of individual elements would result in applied loads being safely distributed to other load carrying members, 90 percent probability with 95 percent confidence.

The use of such precise statistical requirements allows OEMs to make clear goals about the level of confidence that is necessary for their simulation-based results. However, note that both FAA CFR 25.933 and CFR 25.613 make use of probabilistic language that may be inappropriate in the presence of epistemic uncertainty (see Section II).

Current aircraft certification practices already include inherent (and often unquantified) uncertainties due to the nature of test data. In the certification process, this leads to “close or conservative” models, where an acceptable level of mismatch between data, models, and requirements is included. For example, flight test data is known to include some amount of error due to process noise and measurement noise, but the exact amount is unknown, directly leading to some level of conservativeness in performance metrics. For certification by analysis, these differences directly impact the level of confidence required in the analysis to match or improve the level of safety and accuracy when compared with testing based validation.

The last impediment to overcome within this category is the difficulty of engineers and analysts in communicating uncertain results to decision makers and certification authorities. This challenge, in part, could be made less severe over time with the aforementioned suggestion to increase statistical education in undergraduate engineering programs. In the nearer term, communication of uncertain results might become more effective if the industry or CbA community were to agree upon a consistent set of terms, types of plots, and types of UQ analysis to be used for certification purposes. Reducing the entire possible set of UQ analyses and presentations into a smaller suggested subset with straightforward interpretations would help non-UQ-experts to better-understand uncertainty analysis in the context of CbA.

III.B. Tools & Capabilities

Even if the Mindset and Awareness impediments in Section III.A are addressed, there are still many technical challenges for UQ tools and capabilities that must be overcome to allow for expanded use of CbA. One of the most important items in this category will be the development of methods to rigorously and efficiently propagate both aleatory (inherent, probabilistic) and epistemic (non-probabilistic, lack of knowledge) uncertainty. Ferson and Ginzburg¹² were early advocates of segregated uncertainty propagation approaches for mixed (aleatory + epistemic) uncertainty, but progress in this area has stalled in recent decades. Moreover, many organizations rely heavily on Bayesian methods for uncertainty analysis, forcing practitioners to improperly assign uniform prior distributions to sources of uncertainty that rigorously should be treated as non-probabilistic epistemic intervals or p-boxes. The expansion of or departure from Bayesian methods to include epistemic uncertainty will be a difficult technical and organizational challenge.

Epistemic uncertainty is also important for quantifying model form uncertainty. This is a type of uncertainty that arises due to the inability of the governing equations, mathematical formulations, and other conceptualizations and abstractions of a model (i.e., its model form) to describe the physics of the real device or system being simulated. It is fundamentally due to lack of knowledge, and should be treated as epistemic. Estimates of model form uncertainty are a product of model validation, during which simulation results are compared to physical test results with uncertainty considered in each. Approaches such as area validation metrics⁹ or real-space validation¹³ have been proposed to quantify this statistical difference, but more work

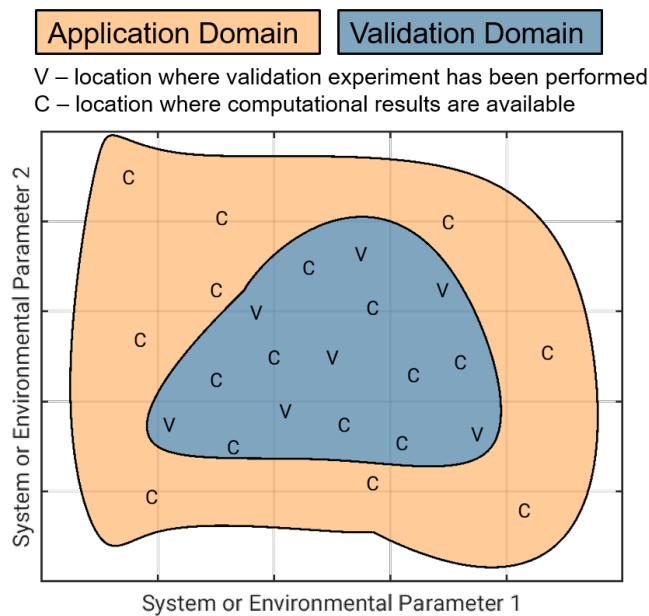


Figure 3. Depiction of a validation domain and an application domain; estimating model form uncertainty in the latter is especially difficult.

is needed to understand the limitations of credibility assessments which can be made from validation and model form uncertainty estimates. Of particular relevance to CbA is the idea of interpolating or extrapolating model form uncertainty from the validation domain, where model validation has been performed, to the application domain, where no physical test results are available (see Fig. 3). For example, an OEM may perform validation of CFD results by comparing to flight test data at a select set of Mach number and angle of attack conditions, but may wish to assess the credibility of the same CFD process to make predictions in a different region of the flight envelope where no test data is available. Another example would be to transfer model form uncertainty from an existing physical design to a new design that incorporates different technologies and configurations. New approaches are needed to perform this type of analysis, and must include statistically defensible credibility assessments which can be agreed upon between OEMs and regulators.

Given the complexity of a commercial aircraft, further development is also necessary to enable UQ for multi-disciplinary, multi-physics analysis. Fluid-structure interactions, ice accretion modeling, and many other types of coupled physics analysis can be difficult to perform even deterministically, and yet the implications of their results may be paramount to aircraft safety and certification. If these types of analyses are to be used for CbA, then methods must be developed to quantify parametric and non-parametric sources of uncertainty in the individual physics domains as well as in their interacting terms or processes (e.g., the interpolation of flight loads from a CFD solution to a FEM mesh). The CbA community should work together with the Multi-disciplinary Design, Analysis, and Optimization (MDAO) community to address these issues. The MDAO community has experience developing robust design and design under uncertainty methodologies, and the same capabilities which have been developed for multi-disciplinary design spaces may also be applicable to multi-physics flight modeling in the context of CbA. Moreover, MDAO capabilities could be leveraged or adapted for the interpolation and extrapolation of uncertainty throughout a flight envelope.

A special focus is needed for efficient UQ methods when performing the aforementioned tasks because in most cases the applied deterministic analysis tools are already resource and time intensive. Hence, methods that require thousands or more model evaluations are simply infeasible. Note that, when deriving efficient UQ propagation techniques, the focus should always be on the scalability and direct applicability for industrial-grade problems. A common approach is employing surrogate and/or reduced order models that aim to emulate the behaviour of the costly model. However, such an approach introduces another source of uncertainty (surrogate model error) that must be considered.

None of the capabilities discussed in this section will be useful unless they are made available to engineers and analysts. For this reason, software tools must be developed – either internally by OEMs, externally by software vendors, or in collaboration with government partners – with CbA in mind. These tools should

Table 2. Completeness Levels of experimental attributes, as described by Oberkampff and Smith.¹⁵

Completeness Level	Description
Level 0	Little or no information is provided on an attribute.
Level 1	Some information is provided on an attribute, including assumptions made and an estimate of total experimental uncertainty.
Level 2	Detailed information is provided on an attribute along with a discussion of assumptions made and an estimate of total experimental uncertainty.
Level 3	Detailed information is provided on an attribute (e.g., highly resolved spatial and temporal measurements), along with an extensive discussion of assumptions made and data processing procedures, and a comprehensive estimate of total experimental uncertainty.

be well-documented, and must not assume that their users have a strong background in statistics (see Section III.A). They should include a range of UQ methods, along with recommendations of which methods are most appropriate based on factors such as the computational cost of function evaluations and required accuracy for the statistical results (for example, surrogate modeling methods may be appropriate for estimating the bulk magnitude of uncertainty, but may be less accurate for assessing low-probability safety-related events in the tails of distributions). Given the safety implications of their end-use, the delivery of these UQ tools should also include code and solution verification documentation to ensure that the statistical methods employed are free from coding errors. The authors of this document perceive the current state of UQ to be full of small teams writing their own codes for research purposes; this siloed approach is important for the development of new methods, but it prevents the democratization of UQ that could occur with a set of dedicated tools and resources for CbA UQ efforts.

III.C. Data & Benchmarks for V&V

Verification & Validation (V&V) is a fundamental step towards CbA and addresses analysis tools as well as UQ processes. To be fair, it is a well-established step in every analysis tool development, or at least claimed to be so. However, in most cases a deterministic approach is pursued in which analysis results are compared to reference data without accounting for uncertainties. Instead of deriving credibility estimates, typically binary-like statements are provided. As an example, a turbulence model is labeled either useful or not for a specific flow phenomena of interest.

As V&V heavily relies on data against which results can be compared, having high-quality validation experiments at hand is crucial. Even though these types of experiments have been performed at various institutions throughout the years, nearly all available data sets lack a rigorous identification and characterization of uncertainties. The authors identified a particular lack of knowledge with respect to epistemic uncertainties present in validation data sets. The authors also note that many experiments do not follow the guidelines for a proper validation experiment (as described in Refs. [14, 15]) but rather aim at gaining large amounts of deterministic data with little to no experimental uncertainty quantification. Unfortunately, resulting data sets have limited usefulness for improving the physical accuracy or assessing the credibility of codes. Future experiments should be conducted with the purpose of advancing V&V, UQ, and CbA goals, and experimentalists should strive to attain high Completeness Levels for experimental attributes such as boundary conditions and fluid properties. A description of each Completeness Level is included in Table 2; see Oberkampff and Smith¹⁵ for a detailed description of six independent experimental attributes and scoring of their levels.

Another issue with currently available experiments is that they are normally tailored towards a specific physical phenomena and hence lack the complexity of large scale, industrial-grade cases which are needed in order to advance the state of the art with respect to simulation model V&V and UQ. When addressing these benchmark cases, enforcing a blind test character instead of allowing explicit model calibrations due to known results has been found to be beneficial for building confidence in the proposed methods and tools.

As outlined before, multi-disciplinary, multi-physics analyses are necessary to simulate modern commercial aircraft. Hence, also V&V is needed for these type of complex simulations calling for appropriate data to rely on. A particular focus should again be on large scale, industrial-grade benchmark cases that are made available within the community. While this has been partially achieved through various AIAA interest groups and workshops (e.g., the NASA CRM configuration¹⁶), available benchmarks should ideally also include flight test data. This poses a significant challenge because highly complex and expensive experiments such as flight tests are normally not shared within the community, but instead kept confidential within the few OEMs or large-scale research institutions that can afford to perform them.

When considering the application of UQ to CbA the list of the relevant uncertainties in simulation models is directly linked to the route to certification and the confidence of demonstrating compliance with the requirements relative to established aircraft certification practices. The established certification practices include many “acceptable means of compliance” which address different aspects of the potential model uncertainties. As an example from EASA 25.471(b):

EASA 25.471(b). Critical centres of gravity. The critical centres of gravity within the range for which certification is requested must be selected so that the maximum design loads are obtained in each landing gear element. Fore and aft, vertical, and lateral aeroplane centres of gravity must be considered. Lateral displacements of the centre of gravity from the aeroplane centreline which would result in *main gear loads not greater than 103%* of the critical design load for symmetrical loading conditions may be selected *without considering the effects* of these lateral centre of gravity displacements on the loading of the main gear elements, or on the aeroplane structure provided

- (1) The lateral displacement of the centre of gravity results from random passenger or cargo disposition within the fuselage or from random unsymmetrical fuel loading or fuel usage; and
- (2) Appropriate loading instructions for random disposable loads are included under the provisions of CS 25.1583(c)(1) to ensure that the lateral displacement of the centre of gravity is maintained within these limits.

In this case, the impact of uncertainties in variations in the lateral center of gravity (CG) on the aircraft design loads do not need to be considered if their impact is less than 103% of the critical design load. The resulting uncertainty in the design loads envelopes due to lateral CGs are therefore accounted for (although not explicitly quantified) within the corresponding certification safety factors. This is one of many examples of acceptable means of compliance to the certification requirements for aircraft with different levels of uncertainties resulting in the same physical design. This highlights an inherent (but unquantified) variation in the acceptable level of confidence in the certification requirements.

III.D. Applied Research & Established Processes

Analysis tools are already routinely used for complex, industrial-grade cases and showcase exceptional results. Moreover, common use-cases have been defined within the community that are regularly used to compare results with respect to accuracy and run-time. Good examples are the AIAA drag prediction and high-lift prediction workshop series.^{17–19} For both, the blind-testing nature has helped to advance capabilities and shape best-practices. Various research activities are also ongoing in the UQ field. However, there is a coexistence rather than a combination of analysis tools and UQ methods. In fact, applied research for UQ methods with a clear focus on relevant cases using complex analysis tools is missing nearly entirely.

In order to advance UQ methods and establish their strengths and weaknesses, the applications which are used for demonstration need to evolve towards large-scale, industrial-relevant cases as illustrated in Fig. 4. Similar to other AIAA activities, a set of shared industrial-grade use cases would be highly beneficial for guiding the community towards this level of fidelity; moreover, the selection of these cases could be informed by the near-term and long-term target CbA applications suggested in the NASA Guide on CbA.⁴ Initial

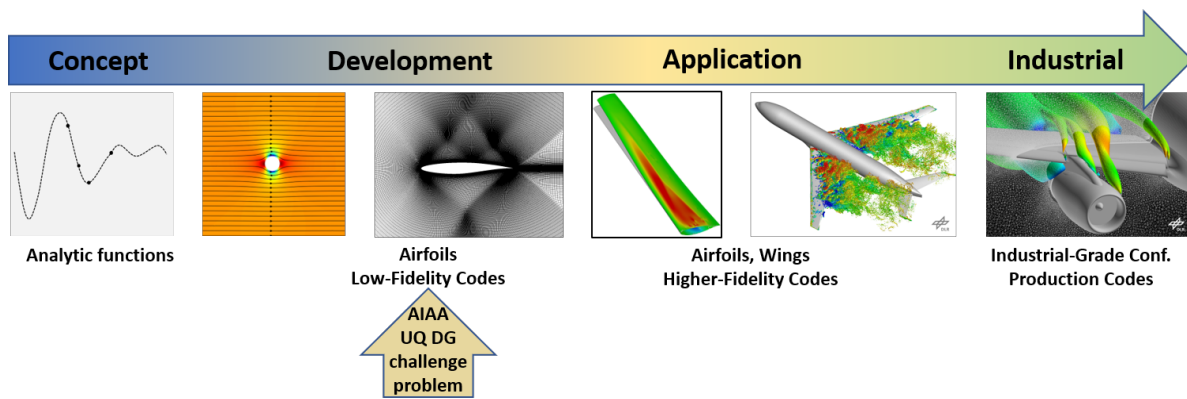


Figure 4. Spectrum of UQ application case complexity towards industrial-grade.

steps have been taken by the AIAA Fluid Dynamics TC UQ Discussion Group to define a UQ challenge problem^{20,21} which allows the community to apply and compare UQ methods on an aerodynamically meaningful use-case. However, the complexity of the underlying problem – aerodynamics based on linear potential theory – does not compare to established analysis tools which are envisioned to be used for CbA such as CFD or even multi-disciplinary, multi-physics analysis. While the UQ challenge problem is very focused on rigorously propagating and quantifying uncertainties and deriving statements about credibility, a large amount of implicit knowledge (often times labelled as “expert knowledge”) is present within different organizations. Currently there is no clear path visible on how to incorporate this knowledge in the context of UQ for CbA.

Ultimately, UQ has to become an established process either as part of existing tools or as additional tools that allow engineers, who are not UQ experts, to perform UQ analysis and obtain the results required for CbA. The authors envision the future usage of UQ to be similar to the evolution of CFD usage, which in its current state can be seen as a readily available tool that engineers, academic professionals, and students are all able to use to gain aerodynamic insight. Significant advances in this area are necessary and involve applied researchers to derive best-practice guidelines as well as software providers to either integrate UQ capabilities within their existing tools or to provide specific UQ tools. Closely related to the aforementioned point is the question of best-practice guidelines that are applicable to the analysis tools at hand. This encloses the four steps of the UQ process presented in Section II, which are identify, characterize, propagate, and analyze uncertainty. Again, a shared set of use-cases could help to answer several of the questions and allow to derive a commonly agreed-upon UQ best-practices document.

IV. Conclusions

Aircraft certification by analysis promises to enable reduced development time and lower cost commercial aircraft than today’s modern fleets, while maintaining an equivalent level of safety as the current flight-based certification process. Moreover, CbA has the potential to reduce the risk associated to new technologies that yield a performance increase and a positive climate impact by ensuring that such technologies can be certified in a secure and timely manner. However, significant technical and programmatic impediments must be overcome in the field of uncertainty quantification in order to fully realize the goals of CbA. These challenges can be broadly sorted into four categories: (1) UQ Mindset & Awareness; (2) Tools & Capabilities; (3) Data & Benchmarks for Verification & Validation; and (4) Applied Research & Established Processes. Key enablers for CbA include broader UQ education at the collegiate and professional level, as well as a sharpened focus of existing UQ researchers on applied, industrial-grade uncertainty problems with relevance to CbA.

The path to CbA may include periods where CbA is done alongside traditional testing, periods where CbA is the primary tool but with supplemental testing, and periods where CbA has fully replaced traditional testing. The progress of CbA will also be dependent on the specific discipline – some disciplines or specific tests may lend themselves to CbA, while other fields may not be able to be fully replaced due to the physics or other complexities. Particularly when addressing safety-critical quantities, the transition from factors of

safety to uncertainty estimates must be carefully planned and implemented, with agreement from all relevant communities. Therefore, while UQ can be incorporated at numerous points throughout the CbA process, it should be regarded as a tool that enables the extended applicability of CbA, but not a stand-alone solution. Ultimately, what is sufficient for CbA will be defined by regulatory agencies with input from industry and academia, so clear and continual communication will be key. While these details are being addressed, AIAA and communities therein should take a leadership role in organizing industry, government, and academic partners towards the development of a UQ best-practices document targeted on CbA.

References

- ¹American Institute of Aeronautics and Astronautics, “Recommended Practice: When Flight Modelling Is Used to Reduce Flight Testing Supporting Aircraft Certification,” AIAA Recommended Practice R-154-2021, April 2021, <https://doi.org/10.2514/4.106231.001>.
- ²National Research Council, *Evaluation of Quantification of Margins and Uncertainties Methodology for Assessing and Certifying the Reliability of the Nuclear Stockpile*, National Academies Press, Washington, D.C., March 2009. <https://doi.org/10.17226/12531>.
- ³U.S. Food & Drug Administration, “Assessing the Credibility of Computational Modeling and Simulation in Medical Device Submissions,” Draft Guidance for Industry and Food and Drug Administration Staff, December 2021.
- ⁴Mauery, T., Alonso, S., Cary, A., Lee, V., Malecki, R., Mavriplis, D., Medic, G., Schaefer, J., and Slotnick, J., “A Guide for Aircraft Certification by Analysis,” NASA CR-20210015404, May 2021, <https://ntrs.nasa.gov/citations/20210015404>.
- ⁵Lee, H. B., Ghia, U., Bayyuk, S., Oberkampf, W. L., Roy, C. J., Benek, J. A., Rumsey, C. R., Powers, J. M., Bush, R. H., and Mani, M., “Development and Use of Engineering Standards for Computational Fluid Dynamics for Complex Aerospace Systems,” AIAA Paper 2016-3811, June 2016.
- ⁶National Aeronautics and Space Administration, Standard for Models and Simulations, NASA-STD-7009, 2008.
- ⁷American Society of Mechanical Engineers (ASME), *Guide for Verification and Validation in Computational Solid Mechanics*, ASME V&V 10-2006.
- ⁸Schaefer, J. A., Romero, V. J., Schafer S. R., Leyde, B., and Denham, C., L., “Approaches for Quantifying Uncertainties in Computational Modeling for Aerospace Applications,” AIAA Paper 2020-1520, January 2020, <https://doi.org/10.2514/6.2020-1520>.
- ⁹Oberkampf, W. L., and Roy, C. J., *Verification and Validation in Scientific Computing*, Cambridge University Press, New York, 2010.
- ¹⁰RoCS - Rotorcraft Certification by Simulation, “Preliminary Guidelines for the Rotorcraft Certification by Simulation Process - Draft,” Report of the RoCS - Rotorcraft Certification by Simulation - Project, 2022.
- ¹¹U.S. News, “Best Undergraduate Aerospace / Aeronautical / Astronautical Engineering Programs,” <https://www.usnews.com/best-colleges/rankings/engineering-doctorate-aerospace-aeronautical-astronautical> [accessed March 2023].
- ¹²Ferson, S., and Ginzburg, L. R., “Different Methods are Needed to Propagate Ignorance and Variability,” *Reliability Engineering and System Safety*, Vol. 54, 1996, pp. 133-144, [https://doi.org/10.1016/S0951-8320\(96\)00071-3](https://doi.org/10.1016/S0951-8320(96)00071-3).
- ¹³Romero, V., “Real-Space Model Validation and Predictor-Corrector Extrapolation applied to the Sandia Cantilever Beam End-to-End UQ Problem,” AIAA Paper 2019-1488, 21st AIAA Non-Deterministic Approaches Conference, AIAA SciTech 2019, Jan. 7-11, San Diego, CA.
- ¹⁴Aeschliman, D. P., and Oberkampf, W. L., “Experimental Methodology for Computational Fluid Dynamics Code Validation,” *AIAA Journal*, Vol. 36, No. 5, pp. 733-741, May 1998, <https://doi.org/10.2514/2.461>.
- ¹⁵Oberkampf, W. L. and Smith, B. L., “Assessment Criteria for Computational Fluid Dynamics Model Validation Experiments,” *ASME Journal of Verification, Validation, And Uncertainty Quantification*, Vol. 2, No. 3, September 2017, <https://doi.org/10.1115/1.4037887>.
- ¹⁶“NASA Common Research Model,” <https://commonresearchmodel.larc.nasa.gov/> [retrieved November 2023].
- ¹⁷“Announcing the 6th AIAA CFD Drag Prediction Workshop,” <http://aiaa-dpw.larc.nasa.gov/> [retrieved November 2023].
- ¹⁸Vassberg, J. C., Tinoco, E. N., Mani, M., Rider, B., Zickuhr, T., Levy, D. W., Brodersen, O. P., Eisfeld, B., Crippa, S., Wahls, R. A., Morrison, J. H., Mavriplis, D. J., Murayama, M., “Summary of the Fourth AIAA CFD Drag Prediction Workshop,” AIAA Paper 2010-4547, June-July 2010, <https://doi.org/10.2514/6.2010-4547>.
- ¹⁹“The 3rd AIAA CFD High Lift Prediction Workshop (HiLiftPW-3),” <https://hiliftpw.larc.nasa.gov/index.html> [retrieved November 2023].
- ²⁰Cary, A. W., Schaefer, J. A., Duque, E. P. D., Khurana, M. S., and DeCarlo, E. C., “Overview of Challenges in Performing Uncertainty Quantification for Fluids Engineering Problems,” AIAA Paper 2022-2357, January 2022, <https://doi.org/10.2514/6.2022-2357>.
- ²¹AIAA Engage, “Uncertainty Quantification Fluid Dynamics TC Discussion Group,” <https://engage.aiaa.org> [accessed March 2023].