Twins, pyramids and environments: unifying approaches to virtual testing

Ms L. Wright, Dr. K. Khatry, Dr. M. Bevilacqua, Dr. L. Wright, Dr. M. Chrubasik, Dr. J. Gregorio
(National Physical Laboratory, UK);

Abstract

There is a long history of the use of engineering simulation for design. This virtual approach is typically followed by physical prototyping, testing and refinement to reach a final design, followed in many cases by a physical testing regime to meet regulatory requirements. For many companies, the use of simulation tools has reduced the time and cost associated with getting new products to market due to the ability to explore multiple designs, and has reduced resource usage and improved product quality by enabling exploration of aspects of manufacturability and long term in-use performance.

Companies are increasingly seeking to gain similar benefits beyond the design stage. Some companies whose products are subject to extensive regulatory testing requirements are seeking to provide evidence of compliance through a combination of simulation and testing. It is common in such industry sectors to have a "testing pyramid", where the safety of a complex multi-component product is demonstrated by carrying out tests of materials, components, assemblies and complete systems, with the number of tests carried out decreasing as the complexity of the object under test increases. A "smarter testing" approach would replace some of these physical tests with simulations and would feed information between the various tests to improve the validation of the simulation and the confidence in the evidence of safety.

Some products cannot be fully tested via physical testing alone because they cause a risk to human safety. An example of current relevance is autonomous vehicles. The artificial intelligence (AI) that controls an autonomous vehicle is often trained on data obtained from human-controlled journeys of a vehicle with the sensor suite in operation, so that the AI is shown what safe driving looks like under typical conditions. However, many of the situations most likely to lead to an accident are not encountered in typical driving conditions, and could cause risk to life if recreated deliberately. A simulation can potentially recreate high-risk scenarios safely for both training and testing purposes.

Some products can significantly improve lifetime prediction and understanding of real-world performance by linking models and data in a digital twin. This approach can lead to improved future design iterations and more effective maintenance plans as the understanding of the product improves. Application of this approach could support personalisation of devices such as medical protheses, where monitoring, adjustment and individualisation could significantly improve people's lives.

In all of these applications it is important to note that staff at the company developing the simulations are not the only people that need to have trust in the simulation results. That trust needs to be shared by regulators, end users, and in some cases the general public.

These three seemingly distinct themes of activity are strongly linked, not least because they have the same need underpinning them: they need to combine validated models and measured data to make trustworthy predictions of real-world behaviour. This need can be answered most efficiently by a combination of activities in several technical areas, including data quality assessment, software interoperability, semantic technologies, model validation, and uncertainty quantification. The technology readiness level of these areas is varied, and the level of awareness and uptake of good practice of each technical area varies across sectors.

This paper discusses the common features, and differences between, the fields of smart testing, virtual test environments, and digital twins. Starting from a consideration of commonality we will highlight areas where existing methods and expertise could be better exploited, and identify areas where further research and development of tools would accelerate successful application of trustworthy digital assurance approaches in industry.

1. Introduction

The industrial usage of computational engineering tools has evolved hugely since NAFEMS was first founded. Developments in computer hardware and solver technology, along with advances in adaptive meshing and contact algorithms, have broadened the range of problems that are now addressed using simulation technology. Manufacturing processes can be developed using dynamic multi-physics models to identify optimal operating conditions. Structural integrity and product performance can be assessed over long timescales, enabling manufacturers to understand behaviour of devices over their full lifetime and to take a more informed approach to maintenance for key assets.

This wider use of simulation technology has led manufacturers and industrial users to look for other areas where simulation might be able to save them time and money. One area that has been identified by several industries as being of potential benefit is the use of simulation results within regulatory approval processes, often known as "certification by analysis". In many sectors, regulatory approval requires provision of physical test results (typically generated by following technical standards) to demonstrate that a product is safe for use and should be approved. For some sectors, such as aerospace, the physical tests required include tests of materials and sub-structures that are critical to the safe operation of the product as well as full-scale product testing. Many of these tests are tests to destruction, and in most cases they must be repeated several times to demonstrate the variability of the results. The process is therefore costly in terms of time, money, and resources. The use of simulation would potentially be quicker, cheaper, and less wasteful of material.

From the regulator's point of view, the physical testing approach generates reliable evidence that a product will perform as it is supposed to and will not cause harm when used. The evidence required by the regulators has been developed based on a deep understanding of the operating conditions and likely failure modes of the products being regulated, so the regulator can have confidence that the risks associated with use of a product have been assessed. The use of standardised tests gives the regulator confidence that the testing has followed good practice, and the use of repeated tests gives confidence that the passing of a given test was not a chance occurrence. The use of simulation results in regulation will require similar mechanisms for confidence to be created so that regulators can be sure they are continuing to fulfil their remit.

For some regulators, consideration of simulation results has become more pressing because autonomous systems are being developed for areas that fall within their remit [1, 2]. An obvious example is autonomous vehicles, where systems for road, marine, and air deployment are being developed. These systems cannot rely on the "common sense" that we assume that human operators will use when encountering new or unusual situations, so the testing of these systems must be more extensive, and testing must include rare but high-risk situations. The importance of such situations, which the system would not be exposed to during normal driving and which are too dangerous to recreate in real life, to human safety means that virtual testing, and hence simulation, will form an essential part of regulation of autonomous vehicles. Whilst these events are rare, the large number of road vehicles in use means that they will still occur and thus risk causing harm.

Certification by analysis (CbA) and virtual testing are both situations that require a verified and validated model of the real world, and therefore have much in common with digital twins. Digital twins [3] combine measured data gathered from a real-world object or system and a validated model of that system to produce a model that can be regularly updated to give an accurate reflection of the current state of the system and can therefore be used to predict future behaviour and test possible scenarios using the current state as a starting point. Potential applications include predictive maintenance, active traffic flow management, and tailored management of high-value assets.

These three applications of engineering simulation have much in common in terms of their technical needs, and development in the three areas could be accelerated by sharing of good practice. This paper highlights the areas of commonality and the distinctions between the areas, specifies requirements in the areas of commonality, and discusses technologies and recent work to address those needs.

2. What makes these applications special?

At this point it is worth considering why CbA, virtual testing and digital twins are different from the applications of many engineering models. It is important to consider that all engineering applications of model results need confidence at some level, so the need to have trustworthy results is not unique to these applications.

One key consideration is that these approaches are used when a product or system is at or very close to the deployment and usage stage. If the simulated products are unsafe or do not perform the way they are supposed to, there will be physical harm or financial loss for the end users and a resulting loss of reputation and potentially legal repercussions for the manufacturer. The level of confidence that is required in the model results is therefore higher than for models used in a design process, when products can still be adjusted before deployment, or for lifetime prediction where other checks on safety are likely to be in place during deployment.

A further complication for CbA and virtual testing in particular is that the models will almost certainly be created by the manufacturer of the product but they need to be trusted and accepted by the regulator. Regulators are unlikely to have the capacity and the software access to be able to investigate every model used in CbA and virtual testing themselves. Instead they will place the burden of proof of model adequacy on product manufacturers by requiring appropriate evidence of validation, uncertainty quantification, data quality, model integrity, and other aspects of trustworthiness.

An additional factor for virtual testing and for some applications of certification by analysis and of digital twins is complexity, with the complexity appearing in a different form in each case.

Virtual testing is carried out in a virtual testing environment (VTE). Several tools are already available for road vehicles in particular, generally focussed on photo-realistic simulation and accurate simulation of vehicle motion and physics [4, 5, 6]. The purpose of the VTE is to simulate test scenarios with sufficiently high fidelity that the algorithm under test would produce the same response if it underwent the test in the real world. This purpose requires the VTE to include simulations of anything that might affect the result of the test in the real world. The result of the test is governed by the response of the algorithm under test to its environment, which is determined by the data the algorithm receives from its sensor suite. The VTE must therefore contain sufficiently accurate models of the sensors, including incorporation of any factors that affect sensor performance (for instance rain may scatter lidar signals). This idea is sketched in Figure 1 below: the VTE includes all interactions between any part of the autonomous vehicle and the real world. This requires the VTE to include a way for the sensor models to probe their environment and receive an appropriate response, so for instance it may become necessary to assign radar cross-sections to significant objects in a VTE if a vehicle has a radar sensor. The need to describe complex random effects such as weather and to assign appropriate properties to objects within the domain rapidly lead to a complex system.

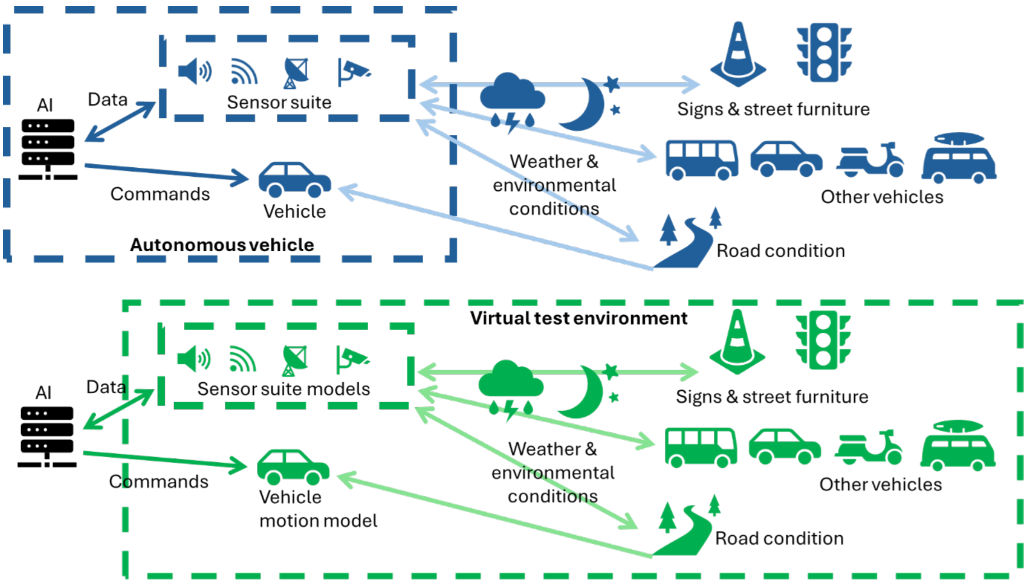

Figure 1: Sketch comparing the interactions between an autonomous vehicle AI and the real world (upper sketch) and between a virtual test environment and an AI (lower sketch).

Certification by analysis will vary in complexity from sector to sector. Sectors such as aerospace have a multi-level "testing pyramid" that provides cumulative evidence of product safety. The testing pyramid is sketched in Figure 2, showing that the following tests are carried out:

- a very large number of tests on materials,

- a large number of tests on structured coupons,
- a medium number of tests on component parts,
- a small number of tests on assemblies and
- a very small number of tests on the complete product.

Information from the lower-level tests can be used to inform the design and verify the results of the higher-level tests. The goal of certification by analysis in aerospace is to replace some of this testing pyramid with analysis, including the propagation from one level to another. The ideal would be to propagate not only results from one level to another, but also to propagate the uncertainty associated with measured test results and with simulation results, in order to build the most complete picture possible of the reliability and safety of the product, more quickly and more efficiently than current processes allow.

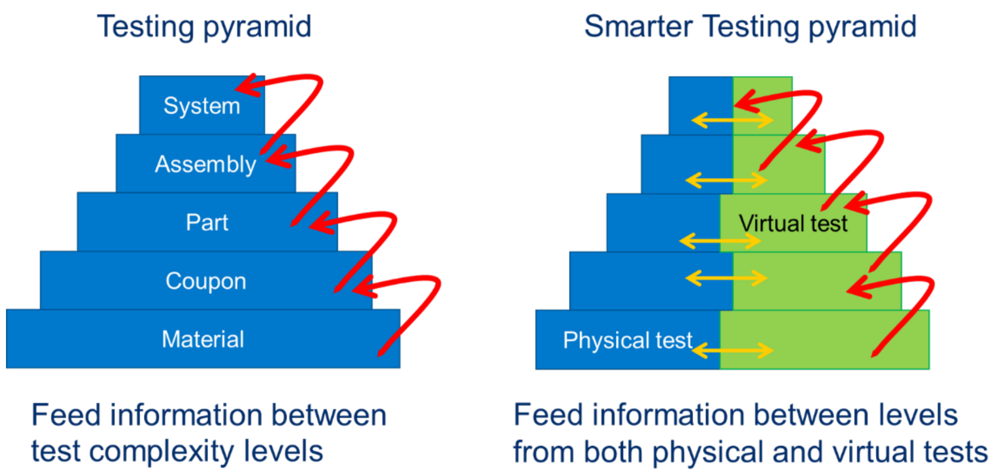

Figure 2: Sketch comparing testing pyramid with the Smarter Testing approach.

This testing pyramid approach could also be used in autonomous vehicles to design the balance and links between virtual and real tests, and also to scope a set of tests covering the sensor suite (including assessing sensitivity to weather) as well as the complete vehicle.

The complexity of a digital twin will depend on the complexity of the physical object being twinned and the factors that affect its behaviour. Twins of a single individual product or object may be straightforward, but twins of large “multi-actor” systems such as cities will be complex. Similarly twins of objects in a stable environment may be simple, but objects such as offshore windfarms that are subject to randomly varying wind and water loading will have complexity that will need to be managed.

3. Needs

These common factors (product maturity, varied stakeholder group, and system complexity) drive the technology needs that are common to all three applications. The system complexity makes demonstration of trustworthiness a complex process; the product maturity means that the level of trust required is high; and the varied stakeholder group means that the evidence supporting trust needs to be communicated clearly.

Trustworthiness of models has been a long-term interest for users of engineering simulation. Initially the focus was on verification (has the selected model been implemented and solved correctly?) and validation (does the selected model accurately describe the real world situation?), known as V&V, and organisations such as the ASME [7] have done much to develop and disseminate good practice guidance on V&V. Many companies that use engineering simulation routinely in their design processes have a formalised V&V process based on the ASME framework. More recently the role of uncertainty quantification (UQ) has been highlighted as important and a VVUQ approach is regarded by many as the gold standard.

It is useful to note that whilst the ASME approach has been extensively applied to validate simulations, there are other forms of complicated calculation arising in certification processes that would benefit from a similar framework. As an example, consider X-ray computed tomography (XCT). XCT is an imaging technique that generates images of cross-sections through an object based on X-ray attenuation. This technique could potentially provide quantitative data on (for instance) internal dimensions of a part to a certification process. However, the process of going from the raw image to a numerical value involves various image processing techniques that impose assumptions about the object or the data and depend on imaging parameters such as pixel size as well as algorithmic parameters such as binarizing thresholds. The process of validating these calculations and choices against a ground truth is an important step in generating trust in the output of the system. In the longer term, standardisation of the methods used to carry out such calculations will ensure that a trustworthy approach is used.

Similarly, the topic of Simulation Data Management (SDM) has been a long-term concern for simulation engineers in many industries, with many software houses offering tools to support effective SDM and a well-established NAFEMS working group (SDMWG) providing guidance and examples. Effective management of simulation results and simulation quality more broadly will enable manufacturers to provide the evidence that regulators and certification bodies need.

The existing good practice in VVUQ and SDM provides a solid foundation on which CbA can build. However, in the case of CbA, VTEs, and digital twins, the complexity of the systems under consideration and the need for communication of trustworthiness to a wider range of people means that additional approaches to manage complexity, quantify confidence, and communicate trustworthiness are essential in order to increase trust and reduce risk. There are several areas that are useful to consider in more detail as tools for risk reduction in the presence of complexity: data quality, metadata and machine readability, semantic technologies, and uncertainty quantification.

At the National Physical Laboratory, the UK's National Metrology Institute, a multi-disciplinary team has been working on tools, frameworks and guidance documents to enable end users to integrate these complexity management ideas into their workflows. The references in the following sections highlight some of our most recent work, carried out in collaboration with industrial end users and other research bodies.

Data quality

Data quality is an umbrella term that encapsulates many different properties of data that may affect how useful it is for informing decision-making. Different applications and different industries prioritise different aspects of data quality. For instance process control systems are likely to need timely data but may be tolerant of incomplete or missing data, whereas design processes may need complete data but are prepared to accept historical or estimated data instead of current data (at least initially). This range of user needs has led to sector-specific de facto standards and, in many cases, inconsistent terminology. This inconsistent terminology has the potential to cause major problems when applications span several technical areas, which is likely to occur for CbA for complex products where tests of many different types of system are necessary.

The importance of data quality has led to development of a range of standards and publications defining data quality in various industries and domains. ISO standard 25012 [8] is focussed on data quality in the context of software development. According to this standard, data quality relates to the extent to which data meets the specifications established by an organisation responsible for developing a product, which again highlights the importance of the end goal which would drive the specifications. This standard defines data quality dimensions as inherent, contextual or system-dependent and provides categories (e.g. precision and accessibility can both be found under contextual) for each group to enable users to identify and consider the dimensions that matter most for their applications. A recent paper [9] used these categories to inform a literature survey of 173 papers about data quality and assigned the various terms found therein to the ISO 25012 categories to try to create a more unified vocabulary; Figure 3 below shows the generated categorisation and emphasises the complexity of "data quality" as a concept. The paper also highlighted two dimensions that were not present in ISO 25012 that may warrant further investigation.

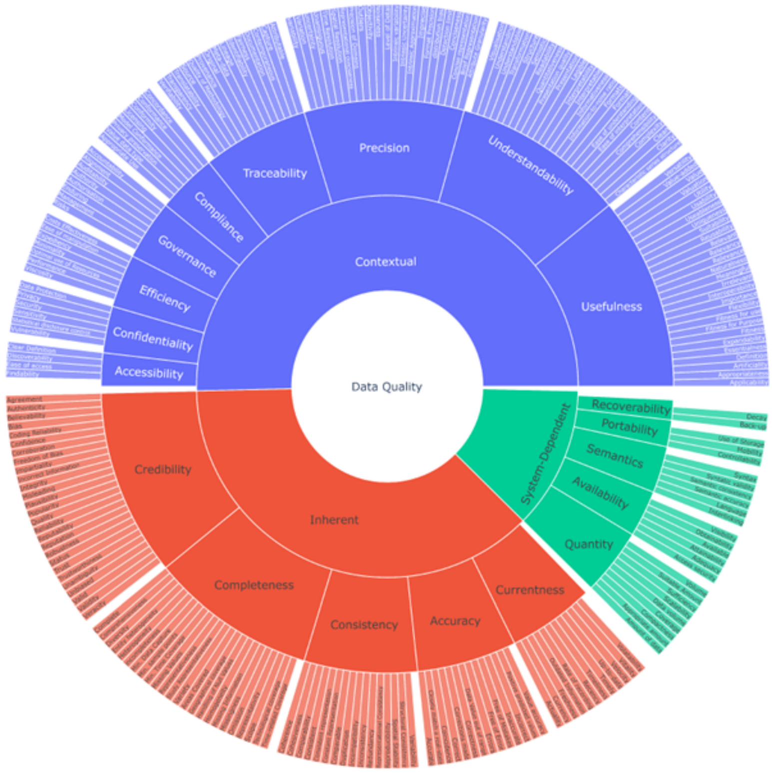

Figure 3: Diagram from [9] to illustrate the complexity and variety of concepts commonly used in literature about data quality.

Related work has led to development of a quality management system for data and software [10] and an associated online interactive risk assessment tool [11] to support development of risk management and quality assurance tools. Whilst the original motivation of these tools was to address management of measurement data quality, the underlying principles are generic and of relevance across sectors.

Metadata & machine readability

Most data is only meaningful with associated contextual information, known as metadata. For example, consider a value of the Young's modulus of steel. The value could depend on the chemical composition of the steel, the method of manufacture, the temperature at which the measurement was made, and various other factors. The trustworthiness of the data might further be affected by whether the test was carried out by a qualified person, whether the measurement equipment had been calibrated recently, the sample preparation method, and so on. All of this information is metadata that needs to be supplied with the data to give the end user confidence that they are using the right values in an appropriate way.

As well as being an important aspect of data confidence, standardised metadata is the first step towards making trustworthy machine-readable data and is the main component of making data Findable, Accessible, Interoperable and Readable (FAIR) [12]. For cases such as VTEs and CbA where data is generated by one party and assessed or used by another, having FAIR data is essential to ensure that evidence supporting confidence in input data is transferred from the model creator to the assessor. One sector that has made significant progress in this standardisation is the civils sector, where standardisation of Building Information Modelling in ISO 19650 [13] provides a framework for creating a shared digital representation of a built asset covering design, construction and operation processes that can be used for making decisions about modifications and maintenance.

Machine readability will require a significant amount of effort to implement. Many public bodies globally hold the type of information that CbA, VTEs and digital twins will need. Examples drawn from the UK include the Ordnance Survey, who hold map data that will inform VTEs for road vehicles, and the British Standards Institution which holds data defining testing standards that will be essential for CbA. Other bodies hold data sets that are of sufficiently high quality to be regarded as reference data, for instance NIST [14] holds various sets of reference data for physical science.

Much of this publicly-held data is held in document form (and may be in proprietary or binary formats), and in many cases whilst the data sources are available as electronic documents (i.e. digitised) they are not available in machine-readable form. Furthermore, in many cases metadata that is important for the intended end use is not available, or the granularity of information does not meet the needs of the end use. As an example, consider a street map. Digitalised street maps contain details of the road network and (to some extent) landmark buildings that would help an autonomous vehicle to navigate, and typically also contain details of road signs and zebra crossings. They do not contain the details of “street furniture”, trees, and non-permanent objects such as roadworks that will help a vehicle navigate over much shorter distances.

Standardisation of vocabulary and a taxonomy of metadata are being advanced in several areas, in particular for autonomous vehicles. This work has involved a collaboration between vehicle developers, regulators, academia, governmental bodies, and independent research institutes. Recent outputs include open pre-normative documents [15-18], national standards [19-21] (typically defining vocabulary and terms at the moment, but the suite of documents will expand in future), and ISO standards [22].

Semantic technologies

Semantic technologies are an important tool for managing complexity and for querying and exploring data. Semantics is the branch of linguistics concerned with meaning, and semantic technologies attempt to link data so that its meaning is preserved. Humans link concepts together through learning and experience, and semantic technologies provide a structure for algorithms to link concepts through defined relationships and logical reasoning in an analogous way.

A semantic approach to structuring data is useful because it provides you with links between data that reflect those that humans make, so the structure can be queried in an intuitive way. The links between the data are ones that humans would construct, so when a human thinks of a question to ask of the data, it can typically be answered using those concepts and connections.

Semantic technologies, which include ontologies, SHACL (shapes constraint language) shapes and knowledge graphs, can also help with, data transformation and integration, validation, interoperability, traceability and lifecycle management of data and objects, and reasoning. They provide an approach to linking relevant, contextual measurement and simulation data and enable better exploration and analysis of that data. They lead to enhanced human-system interactions because they reflect the way humans think about data but also enable machine-understanding of domains.

One of the most widely used semantic techniques is ontology development. An ontology links the concepts within a system together through relations and enables data to be queried by algorithms. Algorithmic queries mean that queries can be automated. Automation of queries brings significant benefit to approval and certification processes [23], since it means that simple tasks such as checking that a particular piece of data was generated by a qualified person within a specified timeframe, can be automated, leaving the qualified assessor to focus on areas requiring human intelligence and potentially reducing the risk of errors being made. Researchers are already investigating how ontologies can support digital twin development and deployment [24, 25], and similar research will support integration in VTEs and within CbA.

Again, much of this is not necessarily new to the field of engineering simulation: best practice in simulation data management implements semantic technologies and automated querying, but in many cases this is achieved through deployment of "black box" commercial tools that may or may not show the underlying data model to the user. Recent work [26] has demonstrated the use of related semantic technologies for automated assessment of simulation credibility, focussing on an example of part reworking. This approach is also applicable to assessment of data quality in its broadest sense.

Uncertainty

Uncertainty is an important approach to capturing the variability that happens in real-world testing. Standards that are commonly used in product testing or material characterisation typically specify the number of times a test should be repeated. This number is often chosen as a balance between the need to get a representative sample of behaviour and the limitations of time and resources. The use of simulation offers an opportunity to carry out more repeat tests (by randomly sampling from the distributions of the model input quantities) and hence gain more insight into the distribution of the model results, because a) once a simulation has been set up there is only a small additional labour cost associated with re-running it, and b) in many cases a simulation can be solved more quickly than a physical test can be carried out.

Uncertainty is also a useful tool for model validation. Validation of a simulation involves demonstrating that the generated solution of the chosen model is a sufficiently accurate representation of the real-world object. This process requires a quantitative comparison of two data sets (the measured and the simulated). Any such comparison process requires a decision to be made about what level of agreement between the data sets is to be regarded as "good enough" for the sets to be considered to be in agreement. Knowledge of the uncertainties associated with the measurements means that the criterion for agreement can be defined in an objective and quantitative way.

Suppose that we wish to compare a measured value of 10 N and a simulated value of 11 N. If we only have numerical values and no uncertainties, we cannot say whether this 10 % difference is significant or not. If we have an uncertainty associated with the measured value then we can define a 95 % (for instance) coverage interval for the quantity of interest and state whether the simulated value lies within that interval. If we have uncertainties associated with both the measured value and the simulated value then we can calculate a measurement of agreement between the two values by evaluating the area of overlap between the two distributions. Confidence in the validation process can therefore be increased if uncertainty information is associated with both the measured value and the model results.

Some simulation engineers are not familiar with uncertainty, so it is important to provide tools and guidance that can be fitted into current workflows and that use examples that are directly relevant to their work. Recent collaborative work in the Smarter Testing project [27] led to the creation of three such tools.

The first tool used hypothesis testing to allow the user to compare simulation results and measured values in a statistically significant way: Figure 4 shows a screenshot of this tool in operation, showing a clear statement of the hypothesis being tested and the result as well as a visualisation of the comparison.

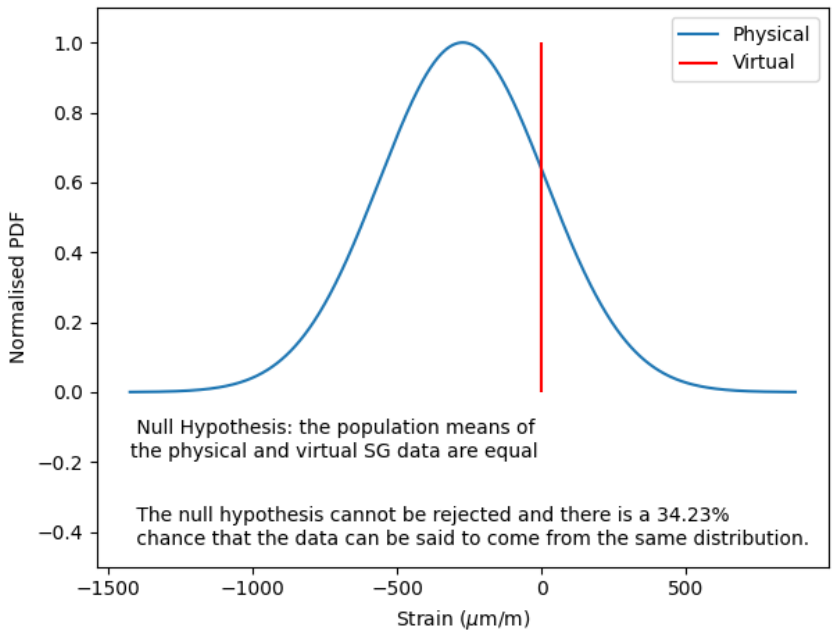

Figure 4: Screenshot of the hypothesis testing tool.

The second tool post-processes model results to evaluate the uncertainty associated with misplacement and misalignment of strain gauges. This information can then be used either to guide gauge placement in tests so that the effects of misalignment are minimised, or to give a better understanding of this component's overall contribution to the uncertainty associated with strain gauge measurement.

The third tool implements some commonly used approaches to sensitivity analysis (Sobol indices [28]) and uncertainty evaluation (random sampling and Latin hypercube sampling [29]) as scripts that interact directly with the Abaqus CAE GUI. The scripts implement a simple test problem (deformation of a beam under an end load and a distributed load) with uncertainties associated with the material properties and the beam cross-section dimensions, and allow the user to explore the effects of changing distributions on the calculated results. These scripts were accompanied by a guidance document that takes the user through all the steps in uncertainty evaluation, from assigning distributions to the input quantities to processing and interpreting the results.

4. Key future challenges

As the sections above show, a lot of progress on underpinning technologies for CbA, VTEs and digital twins has been made. Many of the remaining challenges are around "joining up the dots", making processes more efficient, and delivering usable tools.

A key challenge for the practical deployment of all of these approaches is interoperability. This problem is a long-standing issue in engineering simulation, particularly in industries with complex supply chains. Recent work [30] has looked at the interoperability of Product Management Information supplied by three different CAD packages and has proposed a taxonomy for assessment of interoperability. The work has focussed on a specific application and could be broadened out to other areas. Whilst this approach helps identify data transfers where interoperability may be a problem, it does not make the files interoperable. Interoperability typically occurs when a single robust and well-defined standard is in place and metrics for testing compliance with that standard are in place. At the moment more than one applicable standard is in use, which means that effectively no true standard exists.

Some aspects of the work described above are time- and resource-intensive. In particular, ontology development often involves extensive consultation with human experts to capture their knowledge. Participation in this process means that these experts are taken away from using their expertise. In many cases, some of their knowledge is available in documentary form, in papers, books or standards. Techniques such as Natural Language Processing [31, 32], a form of artificial intelligence, can analyse these documents and obtain the relationships between concepts that feed into an ontology. This type of technology could accelerate ontology development and would reduce the burden on experts and lower the entry barrier to ontology creation.

The section on uncertainty above largely focussed on tools for evaluation of uncertainty for individual models. An essential aspect for CbA in particular is propagation of uncertainty through complex data processing and modelling chains. The transfer of information between and across the levels sketched in Figure 2 will require transfer and use of uncertainty information in order to manage the complexity and preserve trustworthiness.

Finally, it is important to note that whatever work is done needs to be delivered in a form that industrial end users and regulators can integrate into their existing ways of working. There is also a strong need for training to introduce concepts such as ontologies and uncertainty to working simulation engineers so that they understand enough about these technologies to be able to identify when problems are occurring and results are not trustworthy.

5. Conclusions

Simulation offers great opportunities to improve the efficiency of testing, certification, and asset monitoring processes, but the confidence that we already have in physical testing and assessment as a measure of product safety and reliability needs to be recreated in the virtual realm. This confidence is made more difficult to achieve by the complexity of the systems that the simulations need to reflect.

The required confidence can be developed through a combination of data quality assessment, model validation, and uncertainty quantification. The use of semantic technologies and machine-readable data will ensure that the administrative burden on regulators and certification bodies will not be increased and that new knowledge can be inferred from the data that are collected.

A lot of foundational work on these topics has already been carried out, including dissemination of good practice through various NAFEMS working groups. The challenges that remain will build on these foundations, and will require collaborative efforts between domain experts, simulation experts, and data scientists to ensure that relevant technologies are delivered efficiently and in an easily usable way.

6. References

[1] "Connected and Automated Vehicles", Vehicle Certification Agency. Accessed 30/1/25. Available https://www.vehicle-certification-agency.gov.uk/connected-and-automated-vehicles/

[2] "Maritime Autonomous Surface Ships: Creating a framework for efficiency, safety and compliance", Lloyd's Register. Accessed 30/1/25. Available https://www.lr.org/en/knowledge/research-reports/2024/maritime-autonomous-surface-ships/

[3] L. Wright and S. Davidson, "How to tell the difference between a model and a digital twin", Adv. Model. and Simul. in Eng. Sci. 7, 13, 2020. https://doi.org/10.1186/s40323-020-00147-4

[4] "Virtual testing of autonomous vehicles", Claytex. Accessed 30/1/25. Available https://www.claytex.com/wp-content/uploads/2015/01/Virtual-Testing-of-Autonomous-Vehicles_DTC-20161.pdf

[5] "Virtual test drive", Hexagon. Accessed 30/1/25. Available https://hexagon.com/solutions/virtual-test-drive

[6] "Open-source simulator for autonomous driving research", Carla. Accessed 30/1/25. Available https://carla.org/

[7] "Verification, Validation and Uncertainty Quantification (VVUQ)", ASME. Accessed 30/1/25. Available https://www.asme.org/codes-standards/publications-information/verification-validation-uncertainty

[8] "Data quality model", ISO/IEC technical report 25012:2008, 2008.

[9] R. Miller et al, “A Framework for Current and New Data Quality Dimensions: An Overview”, Data , 9(12), 151, 2024. https://doi.org/10.3390/data9120151

[10] K. Lines et al, “A MATHMET Quality Management System for data, software, and guidelines”, Acta IMEKO, vol. 11, no. 4, article 8, 2022. https://doi.org/10.21014/actaimeko.v11i4.1348

[11] “Quality Assurance Tools”, Mathmet. Accessed 30/1/25. Available https://www.euramet.org/european-metrology-networks/mathmet/activities/quality-assurance-tools

[12] M. D. Wilkinson et al, "The FAIR Guiding Principles for scientific data management and stewardship", Scientific Data. 3 (1): 160018, 2016. https://doi.org/10.1038/SDATA.2016.18

[13] “Managing Information with Building Information Modelling (BIM)”, ISO 19650-1:2-018, ISO Standard, 2019.

[14] “Physical Reference Data”, NIST. Accessed 30/1/25. Available https://www.nist.gov/pml/productsservices/physical-reference-data

[15] ASAM OpenODD. Accessed 30/1/25, Available https://www.asam.net/standards/detail/openodd/

[16] ASAM OpenScenario. Accessed 30/1/25, Available https://www.asam.net/standards/detail/openscenario/

[17] ASAM OpenMaterial. Accessed 30/1/25, Available https://www.asam.net/project-detail/asam-openmaterial/

[18] ASAM Test specification. Accessed 30/1/25, Available https://www.asam.net/standards/asam-test-specification/

[19] BSI FLEX 1890 Vocabulary. Accessed 30/1/25, Available http://bsigroup.com/globalassets/localfiles/en-gb/cav/pass-and-flex-pdfs/bsi-flex-1890-v5.0-final-pdf_watermark.pdf

[20] BSI Flex 1889 - Natural Language Description for Automated Driving Systems. Accessed 30/1/25, Available https://www.bsigroup.com/en-GB/insights-and-media/insights/brochures/bsi-flex-1889-natural-language-description-for-automated-driving-systems/

[21] BSI PAS 1881 - Assuring the Operational Safety of Automated Vehicles. Accessed 30/1/25, Available https://www.bsigroup.com/en-GB/insights-and-media/insights/brochures/pas-1881-assuring-the-operational-safety-of-automated-vehicles/

[22] “Road Vehicles — Test scenarios for automated driving systems — Specification for operational design domain”, ISO 34503:2023, ISO standard, 2023. Accessed 30/1/25. Available https://www.iso.org/standard/78952.html

[23] M. Chrubasik, C.T.S. Lorch, and P. M. Duncan, “Ontology-based REST-APIs for measurement terminology: glossaries as a service”, presented at the First International IMEKO TC6 Conference on Metrology and Digital Transformation, Berlin, 2002, http://dx.doi.org/10.21014/tc6-2022.023

[24] K. Inokuchi, J. Nakazato, M. Tsukada and H. Esaki, "Semantic digital twin for interoperability and Comprehensive Management of Data Assets," 2023 IEEE International Conference on Metaverse Computing, Networking and Applications (MetaCom), Kyoto, Japan, 2023,

[25] C. Boje et al, “Towards a semantic Construction Digital Twin: Directions for future research”, Automation in Construction, 114, 2020. https://doi.org/10.1016/j.autcon.2020.103179

[26] J. Gregorio et al, “A competency question driven approach to conceptual data model design for digital verification and validation”, in “Advanced Mathematical and Computational Tools in Metrology and Testing XIII Series on Advances in Mathematics for Applied Sciences, Vol. 94”, F Pavese et al (eds.), pp71-82, 2025. https://doi.org/10.1142/9789819800674_0006

[27] “Smarter Testing”, UKRI. Accessed 30/1/25. Available https://gtr.ukri.org/projects?ref=52036

[28] A. Saltelli, K. Chan, and E. M. Scott (eds), “Sensitivity Analysis”,. John Wiley & Sons, 2001. ISBN 0-471-99892-3.

[29] M. D. Morris, “Factorial sampling plans for preliminary computational experiments”, Technometrics 33(2), 161–174, 1991.

[30] R. Miller, J. Gregorio and P. Duncan, “A framework for streamlining digital supply chain integration through software interoperability analysis”, Measurement: Sensors, 101479, 2024. https://doi.org/10.1016/j.measen.2024.101479

[31] T. Zengeya and J. Vincent Fonou-Dombeu, "A Review of State of the Art Deep Learning Models for Ontology Construction", in IEEE Access, vol. 12, pp. 82354-82383, 2024. https://doi.org/10.1109/ACCESS.2024.3406426

[32] N. Kumar, M. Kumar & M. Singh, “Automated ontology generation from a plain text using statistical and NLP techniques”, Int J Syst Assur Eng Manag, 7 (Suppl 1), 282–293, 2016. https://doi.org/10.1007/s13198-015-0403-1