Simulation Model Quality Issues in Product Engineering: A Review

David J. Murray-Smith, University of Glasgow, United Kingdom, *djms@elec.gla.ac.uk*

SNE Simulation Notes Europe SNE 19(2), 2009, 47-57, doi: 10.11128/sne.19.tn.09939

Dynamic models are of central importance in engineering design for many fields of application but, within some areas, surprisingly little attention is given to the confidence that can be placed on predictions from such models and the implications of model quality, or the lack of it, for design. In recent years the growth of interest in the possible benefits that generic models and model component re-use provide has stimulated new interest in questions of model quality and in the closely associated issues of model testing, verification and validation. This paper considers the link between model quality and the quantitative testing of continuous system simulation models in product engineering and reviews techniques available for the verification and validation of such models. Recent developments and current trends in this field are emphasised, with particular reference to generic models and the re-use of model components. The paper also considers some of the problems inherent in applying rigorous testing and validation procedures. Implications for the education and training of engineering students in the areas of modelling and simulation are considered.

Introduction

In the context of product engineering applications the purpose of a model is to explain complex behaviour, to assist in decision-making processes, or to provide a basis for design. In creating a representation that is appropriate for the intended application there is usually a trade-off between the level of detail included in the model and the speed of solution.

Continuous system dynamic models for the type of product engineering applications under consideration in this paper are most often based on the underlying physics of the system in question but may, to a greater or lesser extent, also involve sub-models that are functional input-output descriptions (i.e. “black-box” models). These may, in turn, be derived from other more detailed physically-based models or may be identified from tests carried out on the corresponding elements of the real system. The models under consideration thus range from completely transparent descriptions based on physical principles, through the intermediate “grey-box” descriptions, to the entirely empirical black-box form of model.

For engineering design applications a good model can have many possible benefits, including early assessment of performance, both within the normal operating envelope and beyond it. Understanding of parameter inter-dependencies and knowledge of key sensitivities within the model can also be of critical importance for design optimisation.

Since a model is, by definition, only an abstraction of the system it represents, perfect accuracy cannot be expected and the key question becomes one of determining the model quality level necessary for the application and assessing the adequacy of a chosen model for some intended use. This implies reducing errors to defined levels for specified regions of the operating envelope of the system. The role of testing, verification and validation procedures can then be regarded as defining boundaries within which a model must operate to specified levels of accuracy. These topics associated with practical issues of model testing are thus of central importance in considering issues of quality in mathematical models and related computer simulations.

As has been pointed out by Sargent (e.g. [1]), Balci (e.g. [2, 3, 4]), Ören [5], Brade [6, 7] and many others, model validation cannot be separated from the model building process. Model building is iterative and, if appropriate methods are used and validation is applied at each stage, confidence in the model should increase from iteration to iteration.

In the early stages of a product engineering design project relatively simple conceptual models are used to examine “what if” situations and allow design trade-off studies to be performed. At this stage little, if any, formal model validation is possible and, inevitably, the error bounds on model predictions are relatively large. Any assessment of model quality and fitness-for-purpose at this point is likely to be based on general design experience and on comparisons with earlier models of other systems having characteristics that are in some way similar. However, as the project moves forward, more complex models may be integrated more fully into the design process and more and more data should become available for model testing. This is likely to involve data at the component level initially, then data resulting from tests on larger blocks and, at a much later stage, data from the testing of complete prototype systems.

Thus, as test data become available, a flow of information starts to be established from the real system to the model, in contrast to the situation at the start of the engineering design process where the flow is entirely from the model to the system being designed. This bi-directional transfer of information is a characteristic of all the later stages of the design and development process. It ensures that the model is being updated continuously as knowledge about the real system is accumulated.

One important development in recent years has been the adoption of a more generic approach to modelling in several engineering fields, including power electronic systems (e.g. [8, 9]) and gas turbine systems (e.g. [10]). In this context the word “generic” is defined as meaning “general” or “not specific” and implies the use of a standard structure and standard building blocks within a model. This approach is likely to become more and more widespread and may be applied in many different application areas in future. The most significant benefit of the generic approach is a more rapid and less costly development process for new models compared with the conventional situation which involves the development, on a one-off basis, of new models for each new design task. Other benefits arise because of the fact that the development and use of a generic model demands a more systematic and rigorous approach to issues of model validation together with better documentation.

Making a model generic in any application area can present difficulties. The essential requirements must be identified first of all and a suitable framework established which provides the necessary flexibility to allow a variety of more detailed needs to be satisfied. Within the generic approach a given system may need to be represented at several different levels of detail at different stages of a design project. This means that sub-models, representing specific parts of the complete physical system, may need to be available at several different levels of complexity, ranging from purely functional forms at the initial stage to highly detailed and fully validated model components for use in the final stages of the project life-cycle. The models at different levels of resolution will, inevitably, all have strengths and shortcomings and need to be mutually calibrated in some way [11]. Ideally, the structures for the different levels of model will be directly related and the models at different resolutions will form an integrated group. The relationship between the different levels of each sub-model within the generic structure must be fully understood by users.

Issues that can arise in the development of a generic model for a new application area have been considered in detail in two recent papers dealing with the modelling of electro-optic sensor systems [12, 13]. The generic model is, in this case, intended to be used in the design of specific types of electro-optic systems such as infra-red search and track systems, missile warning systems and thermal imager system.

The approach adopted for these electro-optic applications involved developing models of specific systems as an integral part of the development of the generic model. Specific configurations of the generic model could then be evaluated and tested, as could modules within the generic description. As confidence in the generic model increased new modules within the generic model structure could be added. However, as the generic model became larger it became more and more important to avoid major changes in the overall structure of the model. Any modifications to a generic model of this kind have to be comprehensively tested using regressive testing methods, similar to those used in software engineering, for particular configurations of the model investigated in earlier tests.

In applying a generic approach to model development, a need may arise to create a model of a new system, not considered already using an available generic structure. This introduces new challenges which encourage re-use of established sub-models but further test the generic philosophy. If the approach fails at any point with a new application then either a flaw has been found in the engineering design or a limitation has been found in the generic model. In the latter case the generic model has to be modified and its capabilities extended.

Modelling errors and uncertainties arise in many different ways, including unjustified modelling assumptions, errors in *a priori* information such as parameter values, inaccuracies in the numerical solution of the model equations and errors in experimental data.

Complex simulation models are sometimes developed and used without rigorous testing and model documentation is often non-existent or inadequate. Poorly tested and undocumented models also may get passed from project to project and thus may end up being used in ways that the original model developer never intended. This contrasts strongly with accepted good practice in the software engineering field where rigorous testing, documentation and version control are all integral elements of the required process for software development. Such methods do not completely eliminate inappropriate or incorrect use of software but they do provide a level of regulation that is often missing in the case of simulation models.

The importance of model quality for product engineering applications was highlighted, about fifteen years ago, in a UK Office of Science and Technology report [14] by the Technology Foresight Panel in the Defence and Aerospace sector. This report includes a statement that "Improved modelling of physical and manufacturing processes will improve our ability to predict the behaviour, costs and risks of future systems and dramatically reduce the development timescale". The report continues with a statement "While it is essential that modelling and simulation is supported by validation trials, improvements will reduce the need for costly and time-consuming developmental testing" [14]. Since that time phrases such as "simulation-based acquisition" and "smart procurement" have entered widespread use within companies involved in defence contracts and have been the focus of discussions within other sectors of industry. In the USA, in particular, the work of the Defense Modeling and Simulation Office (DMSO) within the US Department of Defense (DoD), had significant influence on issues of model testing and of verification, validation and accreditation (VV&A) of models. Although the role of DMSO has been taken over by the Modeling and Simulation Coordination Office (M&SCO) the issues of model quality and VV&A methodology continue to be given priority. M&SCO is involved with annual DoD Modeling and Simulation Conferences and DMSO organised a series of specialist workshops involving staff from government establishments, companies and universities for broad ranging discussions on issues of model quality and techniques for verification and validation (e.g. [15, 16]).

In the USA a SMART initiative (Simulation and Modeling for Acquisition Requirements and Training) has also been established which calls for reuse of models to promote validity, reliability and efficiency of development in areas such as missile systems [17]. The US Office of Naval Research (ONR) has also been very active in promoting new work in this area, especially in the context of power electronic systems and electrical drives. ONR has been responsible for active support of the concept that "the model is the specification" [9]. In other words, it is being suggested that as part of the process of preparing formal specifications for complex new systems a simulation model has to be prepared and that this model becomes the point of reference in determining whether or not the performance of the proposed system is acceptable (e.g. [18, 19]). This means that modelling and simulation activities become a vital element of the acquisition process from the Request for Proposal (RFP) stage onwards. For the customer, the provision of simulation models by competing contractors allows for the comparison of different approaches in a quantitative way at the tendering stage. However, simulations used for such competitive evaluation must have a high degree of transparency and must involve similar sets of assumptions.

US Government laboratories, such as the Los Alamos National Laboratory, the Sandia National Laboratories and the Lawrence Livermore National Laboratory have large research programmes in the general area of model validation methods. Reports on some aspects of the work being undertaken in these programmes may be found in papers presented at the DMSO Foundations `04 V&V Workshop [16].

Within this paper general issues relating to model quality are first reviewed and this leads to a closely-related section of the paper in which methods of verification and validation are outlined. Within a subsection dealing specifically with validation a number of graphical methods are described, together with discussion of quantitative measures and several other approaches to model and system comparison and model analysis. The control systems applications area receives some specific attention. The paper includes a section in which important questions of model documentation are reviewed. This leads to a section involving discussion about the way in which most engineering students are introduced to system modelling concepts within their academic studies and to the inevitable problems if inadequate consideration is given to issues of model verification, validation and documentation at an early stage. The final discussion section attempts to bring together the most important aspects of the review.

1 Model Quality Issues in Product Engineering

There are good examples, often in safety-critical application areas, such as the nuclear industry and the aerospace, defence and marine sectors, where rigorous model testing and formal approval schemes are routinely applied. However, the model development process used within many engineering organisations often involves surprisingly little systematic investigation to establish the quality of the models in terms of their useful range and limits of accuracy. Also, there are many cases where models are justified in a spurious way, possibly on the grounds that the model is one that "has always been used" or is "based on well-established physical principles so must be right" or is "based on an industry standard".

The use of models that are in some ways inadequate for an intended application can often lead to expensive redesign at late stages in the development cycle. The more complex the system being developed, the more likely it is that problems of this kind will arise.

Modelling and simulation activities are important from the concept development stage through requirements analysis to trade-off studies and detailed design. The real system and the associated simulation models generally mature together and the level of model fidelity should increase as a design and development project progresses. Whatever the approach being used for design, experience gained with the real system should be incorporated into the models at every stage.

Many modern developments in engineering involve a "system of systems" design and often require a number of design teams working together. Such collaborative development work means that there is no longer a single "designer" and soundly based, well-understood and well-documented models are essential if all involved in the design effort are to be effective.

Helicopter flight control system design is one example of a field in which model limitations are recognised as a factor that affects the achievable overall performance of the system. Here it is accepted that, until now, the success of modern methods of design has been limited significantly by the accuracy of available models for the vehicle (e.g. [19, 20]). Similar conclusions can be drawn in other application areas in which the eventual performance limits of a new system relate directly to the accuracy of the mathematical model upon which the design is based.

One of the issues that can arise in discussing model quality and validation in the context of control engineering applications is that models used for design are often developed using a combination of physics-based modelling and the experimentally based approaches of system identification and parameter estimation. For example, the structure of the model may be established using physical principles, but values of some of the key parameters of the model may have to be estimated from analysis of results of experiments and tests on the real system. This means that, prior to any experimental work aimed at assessment of model accuracy, a form of testing might have to be carried out as part of the model development process. It is therefore vitally important to ensure that data used in the system identification and parameter estimation stage of model development are not reused at any stage to investigate model quality. However, it is also important that in designing tests for the external validation of such models careful consideration should be given to the range and distribution of the data upon which the identification was based.

It is thus necessary to distinguish carefully between the processes of system identification and parameter estimation that are applied for model development purposes from the processes involved in establishing the quality of the resulting model. The term "model calibration" has therefore been introduced to describe the processes of parameter estimation and other forms of interactive tuning that may be applied to a model during its development. Model calibration is not the same as model validation and these processes take place at different points in the model development cycle.

2 Internal Verification and External Validation of Models

Reasons for errors and uncertainties in models include incorrect assumptions, errors in *a priori* information (e.g. model parameter values), errors in numerical solutions of model equations and errors in experimental procedures and measurements. Much effort has been devoted to trying to separate different aspects of the model development, testing and checking process and to categorise simulation model errors according to their origins [21]. Nevertheless, uncertainty is inevitable since we do not have a complete understanding of the natural world and our measurements and calculations are limited in their accuracy.

An unvalidated model produces results involving unknown and potentially unbounded errors. Even if the user has confidence that the model produces accurate answers most of the time, the situations in which it does not produce accurate output cannot readily be recognised or predicted.

It is important to be precise about the use of words describing the model testing process. It is particularly important to distinguish between the processes of "internal verification" and "external validation". The words "internal verification" describe a process that involves establishing that a computer simulation is consistent with the underlying mathematical model while “external validation” is the process of demonstrating that a mathematical model representing a given real world system is adequate for the intended application [21]. Internal verification is, therefore, the part of the process concerned with establishing whether or not the model is solved correctly, whereas external validation deals with issues of correctness in terms of the structure and parameters of the underlying model description in mathematical and logical terms. This convention is completely consistent with a well-established set of recommendations made in 1979 by the SCS Technical Committee on Model Credibility [22]. Unfortunately, the words “verification” and “validation” are often used in an imprecise fashion. There are also specialist areas (especially in some defence applications such as missile system modelling) where, in the past at least, traditional usage by engineers in some countries reversed the meaning of these two words, compared with the SCS Committee recommendations. It is believed that the inclusion of the adjectives “internal” and “external” helps to reduce the confusion that may otherwise exist when model quality and testing issues are being discussed.

The processes of internal verification of a simulation model are similar to the more general processes of software testing [23] and many of the principles and methods of software testing can be applied. On the other hand, external validation is a more demanding and open-ended task that involves comparisons between the behaviour of the model and the corresponding behaviour of the real system for chosen sets of experimental conditions. This can involve quantitative comparisons of the model’s performance with the real system or a subjective judgement by someone who has a profound understanding of the real system.

Sargent [24] narrowed the definition to emphasise the issue of the accuracy needed for useful model-based predictions in the context of a specific application. This idea can be extended so that external validation is defined as the *confirmation* that the model output has a level of accuracy consistent with the intended use. If this type of approach is used, it is important to ensure that the required accuracy of the model is established prior to the start of the external validation process and not as part of that procedure. Thus, it is often useful to express the results of external validation processes in terms of the appropriateness of the model for a specific application rather than in more absolute terms of a “good” or “bad” description. Indeed, one can never prove that a model is valid; a model can only be proved to be invalid.

For external validation, an important distinction has also to be made between “functional” validation and “physical” validation. The first of these is concerned with the development of a model that mimics the input-output behaviour of the real system whereas physical validation involves establishing the acceptability or otherwise of the underlying assumptions and approximations [25]. As has been pointed out by Hemez [26], perfect matching of all available measured response data is an unrealistic goal and it is more important to ensure that models match available test data with a sufficient level of accuracy for the intended application. This helps to ensure that a given model reproduces test data with an acceptable level of accuracy, while also having a satisfactory robustness to uncertainty. Such uncertainty can be associated with many factors, including modelling assumptions, environmental and model parameter variability or ignorance in terms of initial conditions in the real system. As in control system design, there tends to be a conflict between performance optimality and robustness optimality in modelling [26].

Balci and his colleagues have, in recent years, been stressing the importance of expanding verification and validation from accuracy-centred assessment to assessment which is more quality-centred (e.g. [27]). Quantitative measures of model credibility are hard to define but discussions about the quantification of model credibility may also be found in many sections of the book edited by Cloud and Rainey [28], in the papers by Brade and Köster [29] and Brade, Maguire and Lotz [30] and in the classic textbook on the theory of modelling and simulation by Zeigler, Praehofer and Kim [31].

There have been many suggestions that model testing and accreditation should be more closely linked to ideas of software quality assurance in software engineering (e.g. [29]). This implies improvements in current tools and technologies and also supports the idea that many working in the field of modelling and simulation have much to learn from software engineering principles [23].

2.1 Methods for External Validation

External validation of simulation models is complicated by the fact that most models intended for practical engineering applications involve dozens or even hundreds of quantities that are established and input by the user (e.g. as model parameters), making the problem space very large. Similarly, most models can produce, as outputs, dozens or even hundreds of variables, each of which is likely to contain different levels of error which may also vary with time in the case of a dynamic description. Thus, it is important to establish, *a priori*, which of the output variables of a simulation model are of most interest to the user of the model for the given application. Different users will be interested in different performance measures in different modelling studies and this emphasises the importance of properly matching the model to the intended application at the outset and of establishing *a priori* how much error in the results can be tolerated.

External validation should be considered as an ongoing exercise within the overall modelling process, rather than as a one-off procedure carried out at the end of the model development cycle. It is also important to distinguish between holistic approaches that attempt to validate a complete model externally and model-component approaches in which external validation is carried out at a sub-model level at first. Both are based on the same general principles of external validation but the model component approach may also involve comparisons with test data from component manufacturers.

Confidence in a prediction is a function of the confidence demonstrated in sub-system models as well as in the complete model. This is particularly important where sub-system models can be tested experimentally. Exhaustive testing of sub-system models allows confidence to be established first at the sub-model level and extended gradually to less well-defined situations involving testing of the complete system model over a range of experimental conditions.

In the development of entirely new systems experimental data from the complete system cannot be available at the design stage. In some cases historical data from earlier systems of a similar kind can be helpful in the evaluation of the model of the new system under development. Successful application of this approach depends on good documentation of models of the earlier systems and of the tests carried out to evaluate those system models.

Methods of external validation (i.e. the procedures used to compare observed and simulated values) can be divided into subjective and objective categories. The first approach is based mainly on graphics while the second one involves quantifying the process through specific measures and statistical procedures.

Graphical methods for external validation are typically characterised by plots of simulated values (often continuous and represented by a line) and observed or measured values (usually discrete and represented by points) against an independent variable (often time). One important point of detail, sometimes missed by inexperienced observers, is that the deviation between the simulated and measured values is the vertical difference between corresponding points and should not be assessed simply as the shortest distance between the simulated and measured time history curves.

Another commonly used form of graph involves a simple plot of simulated values against the corresponding measured or observed values. Ideally the plot should be a straight line at an angle of 45 degrees to the axes. Deviations from the ideal are shown by the vertical distance between the points and the 45-degree line and can apply generally to the record as a whole or can be specific to certain sections of the data. Points above the 45-degree line are clearly over-estimated in the simulation while any points below the line are under-estimated. Although viewed by many as subjective, graphical methods are very useful and practical in model validation to complement quantitative measures. Different graphical methods tend to be used in conjunction as different methods of displaying information about a model may provide different types of insight [32].

Quantitative measures for system and model comparison are also very important. The most used deviance measures are the mean-square or mean absolute errors. For the case of n sets of measured and simulated values, the mean absolute error is expressed as the difference between observed values y_i and simulated values $\hat{y}_i$, by:

$$J_1 = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad (1)$$

or using the closely related mean absolute percent error, given by:

$$J_2 = \frac{100}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{|y_i|} \qquad (2)$$

This is a relative error and is inapplicable if any of the observed values happens to equal zero. An obvious disadvantage of these two measures is their sensitivity to single extreme values.

Such an approach can be extended to include some form of weighting function. This means that errors arising in specific sections of the time history can be given special emphasis. One such cost function is:

$$J_3 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^T w_i (y_i - \hat{y}_i) \qquad (3)$$

where w is a weighting factor and the superscript T indicates the transpose.

A measure that has received particular attention for external validation applications in a number of different application areas is Theil's Inequality Coefficient (TIC), which is defined as:

$$J_4 = \frac{\sqrt{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2)}}{\sqrt{\sum_{i=1}^{n} y_i^2} + \sqrt{\sum_{i=1}^{n} \hat{y}_i^2}} \qquad (4)$$

This measure has an advantage in providing values that lie between zero and unity, with values of TIC close to one indicating sets of model and system data that are very different. Values of TIC close to zero indicate small differences between the model and the system time histories.

Other scaled measures are commonly used for comparing model and measured system time histories. Measures based on statistical techniques have received attention in the context of model structure assessment. In particular, step-wise regression and spectral techniques have been used for a variety of practical modelling investigations [21].

One approach, which can be used with benefit in cases where relatively complex models are being considered, involves taking a number of key measured system or sub-system quantities and plotting these as radial lines on an appropriately scaled polar diagram. The length of each line is proportional to the corresponding measure. By constructing a polygon of model measures and a polygon of experimentally determined results from the real system on the same polar diagram an immediate indication of overall model quality is obtained. It should be clear from a comparison of this kind which aspects of the system are represented most accurately and which areas of the model require further investigation. Such diagrams also provide a good way of displaying results from sensitivity analysis of a model. The distortion of the model polygon following a specific imposed change is a useful indication of the overall effect on the model. Polar diagrams of this kind have been used successfully in the context of model testing for electro-optic sensor models [13] and have been considered in the context of fault detection applications as well as in other model testing situations [33]. Although developed independently for the purposes of model test visualisation, these diagrams have many features of Kiviat diagrams which are used in software engineering for visualisation of different metrics associated with software performance and computer hardware evaluation.

All of the quantitative measures mentioned above can also be applied to situations in which one model is being compared with another. This is really a form of verification rather than of external validation. It can arise in situations where a complex, computationally demanding and externally validated simulation model exists but there is a need to derive and test a simpler form of representation which runs on the computer significantly faster. Clearly, the measures and visualisation techniques discussed above can be helpful in the testing and assessment of candidate models in this type of situation, which arises frequently in the development of simulation models that are capable of running in fast timescales, including some real-time applications. An example of this kind may be found in some recently published work of Zenor et al. [34] describing the development of a multi-rate simulation of an underwater vehicle and associated electrical drive system.

2.2 Other Approaches to Model and System Comparison

In some situations, expert opinion plays a vital role in evaluating the suitability or otherwise of a simulation model. For example, a test pilot can quickly establish problem areas in a flight simulator or an experienced plant operator can identify features of a process simulation that do not fit well with his or her knowledge of real process behaviour. In some situations animation can be very helpful in allowing such experts to pinpoint problem areas. Critical examination and correct interpretation of simulation model behaviour from multiple time-history plots is generally far more difficult than viewing the model output in terms of an animation.

Complications arise with methods based on response comparisons when several output variables have to be considered simultaneously or when measurement noise is significant. Methods based on system identification provide a useful alternative to more direct comparisons and can be particularly helpful in giving physical insight about model limitations. The concept of identifiability can also be useful in the design of model validation experiments. Other tools, such as sensitivity analysis, have also been shown to be valuable [21].

Sensitivity analysis can be very important in another way. One very practical approach to external validation (once adequate agreement has been achieved following model calibration activities using system identification tools or other techniques for tuning), involves examining and comparing the effect of changes in the system and the model. For example, in a mechanical system this might simply involve adding mass to some element of the system and changing the corresponding parameter of the system model to test whether the system and model behave in the same way following this modification. If the behaviour is not the same (to some appropriate and predetermined level of agreement) the model will have to be reviewed in terms of its structure and parameters.

Although a generic model can never be fully validated, specific versions of the model can be tested using the general principles of external validation and the measures outlined in Section 2.1 above. More detailed discussion of issues that arise in the testing and external validation of reusable and generic models may be found the work of Malak and Paredis [35]. In the context of automated material handling system design, the paper by Mackulak, Lawrence and Colvin [36] provides useful quantitative information about the benefits of simulation model reuse in terms of model building and analysis for semiconductor material handling applications and provides useful comments on issues of validation in this type of application. Further discussion of the problems inherent in validating generic models may also be found in [12] and [13] for the specific case of electro-optic system models.

3 Engineering Control Systems Applications

Issues of model accuracy have for long been recognised as important in the design of high-performance automatic control systems (e.g. [37, 38]). For high-performance feedback systems it is important to have highly accurate linearised models of the controlled system (the "plant") in the frequency range close to the cross-over region. This is the part of the range where the phase lag for the forward path system transfer function approaches 180 degrees. Model uncertainties within the cross-over region can produce problems in attempting to meet given performance specifications in the closed-loop system.

Much research has been carried out in recent years on frequency-domain modelling for robust control design (e.g. [39]) and on plant model validation by means of system identification methods [40]. However, relatively little consideration has been given to problems of design in highly integrated systems where the traditional division into a "plant" and a "control system" becomes unclear. In particular, we need to consider how we can ensure quality in models that are used for controller design when the plant itself has not yet been completed and is being designed specifically to provide enhanced control capabilities. These are fundamental questions that have already been encountered in the design of advanced aircraft where "control-configured" design has become commonplace. They are likely to have to be addressed in many other control application areas in the future. It is generally accepted that an integrated approach to design should involve the use of generic, externally validated and re-usable sub-models. This is an important issue that is receiving attention in many areas of engineering.

External validation presents particular problems when considered in the context of highly integrated systems. Validation must be iterative and must be carried out in different ways at a number of different stages within the complete design process. With conceptual models at the initial stages of the design process, external validation can only be carried out in a general way. As details of the systems start to evolve validation may necessitate comparisons of reduced models suitable for control system design with computationally more intensive models [41]. At a later stage, detailed testing of sub-systems and hardware-in-the-loop simulation comparisons should become possible. Comparisons may also be made with models that formed the basis of earlier designs of a similar type.

Models are also important for systems that provide automatic fault detection and fault isolation. The critical issue in such systems is to be able to detect faults whenever they occur but avoid false alarms. Fault detection systems that are based on models usually involve monitoring of residuals formed from the differences between corresponding system and model variables. Ideally such residuals are zero in the absence of any fault condition and take non-zero values when a fault occurs. However, non-zero residuals can also arise from measurement noise, unmeasured process disturbances and modelling errors. Appropriate threshold levels for declaration of a fault condition must therefore be chosen. The issue of how to avoid false alarms due to model inadequacies is an important one in such fault detection systems and is closely linked to questions of external validation.

4 Model Documentation

External validation processes do not end when a model is accepted for a particular application. Model documentation, as with documentation of computer software, must allow for changes and further development of the system. Understanding about the limitations of a given model can increase considerably during the application phase of a design project and documentation should be properly updated and maintained for the whole life cycle of the project. This documentation may also be helpful for later developments involving the design of similar systems. Brade [7], as well as emphasising the need for more meaningful documentation and criticising the present lack of quality assurance as an integral part of the model development process, discusses at some length the potential and current limits of documents such as the Verification, Validation and Accreditation Recommended Practices Guide of the US Defense Modeling and Simulation Office [42].

Items in the record for a given model should include the purpose of the model and the intended application, a full model description and the corresponding computer simulation code where applicable, a list of all the assumptions and approximations in the model, details of tests carried out on the real system, details of checks carried out to ensure that the computer-based representation or simulation matches the mathematical description (the process of internal verification) and details of external validation processes applied along with the reasons for accepting or rejecting the model. The documentation should also include statements about the range of applicability of each accepted model.

The process does not end with the decision to accept a model for a particular application. As with the documentation of computer software, the system of model documentation must be capable of accommodating changes and must be updated and maintained for the whole life cycle of the system represented by the model. Regressive testing of models is as important as regressive testing in software projects.

5 Implications for Engineering Education

Methods of model development and testing being applied in industry at present can only be improved if those involved in education recognise the need for change. Engineers are usually introduced to mathematical modelling and encounter computer-based modelling and simulation methods early in their university education. However, the teaching of system modelling methods too often stops with the formulation of equations from physical laws and principles or by system identification and parameter estimation methods. Students are not forced often enough to consider what constitutes a good model and issues of model quality are too often glossed over. Indeed, model evaluation, if considered at all, is often presented as an afterthought rather than as an essential part of the iterative process of model development. Students need to appreciate that correction for model inadequacies can be expensive and time consuming if it is left to the implementation and final testing stage.

In the words of Hardy Cross, a former Professor of Civil Engineering at Yale, "… an important duty of teachers is to force students repeatedly back into the field of reality and, even more, to teach them to force themselves back into reality" [43]. Students must develop an understanding of the limitations of models and for this they need to make critical comparisons of models with real systems. They also need to be required to document models and model testing processes in the same way that they are required to document software that they prepare and test as part of their course-work.

6 Discussion

Validation may be defined as the process of assessing the credibility of a simulation model within its intended domain of use:

1. by establishing whether the simulation model is a correct representation of the underlying mathematical or other formal description (internal verification), and
2. by estimating the degree to which this model is an accurate representation of the real-world system for the intended use (external validation). Whatever the engineering application, the more demanding the system specification the more important it is that adequate consideration be given to these questions that involve issues of model quality.

All models have limitations and the purpose of validation must be to properly define and understand those limitations. However, any practical validation investigation can cover only a finite, and often relatively small, number of test cases. Thus, one should never attempt to prove that a model is correct under all sets of conditions. Instead, a degree of confidence should be established in the model so that its results can be recognised as being reasonable for the objective for which it has been developed. General statements about the validity or quality of a model are therefore inappropriate without reference to its application and the range of conditions considered. One of the inherent problems is the fact that quantitative measures of model credibility are hard to define and as models become more complex there are increasing problems of visualisation.

Continuing research on improved procedures for model development, enhanced computing environments and systematic processes for assessing, correcting and documenting models that are used in engineering design is important. It is also essential that work is directed towards further developing and maintaining libraries of validated simulation models and commonly used sub-models. This is particularly important in terms of being able to fully exploit the benefits of model re-use and the development of generic models.

A strategy is needed to ensure that modelling techniques are properly applied and more effort is needed in all of these areas if we are to reduce development times and costs. The current situation in system modelling contrasts strongly with accepted good practice in the software engineering field where rigorous testing, documentation and version control are an integral part of the recommended processes of software development.

Ideally, what we need is some way of producing confidence intervals for model predictions. Although this goal may be elusive in the case of general nonlinear physics-based parametric simulation models, it is interesting to note that in the Gaussian Process (e.g. [44]) type of nonlinear non-parametric model such additional information is readily available. Also, for linear models, the use of coherence estimates within frequency-domain descriptions of system outputs allows determination of the range of frequencies over which the linear model is applicable (e.g. [20]). More research aimed at applying such techniques to practical engineering problems and developing better ways for assessing the accuracy of predictions from nonlinear physics-based models is essential.

References

[1] R.G. Sargent:, *Validation of simulation models.* In: H.J. Highland, H.J, eds., Proc. 1979 Winter Simulation Conf., Vol. 2, IEEE Press, USA, 1979, pp. 497-503.

[2] O. Balci. *Principles of simulation model validation*, Transactions of the Society for Computer Simulation International, Vol. 14, No. 1, 1997, pp 3-12.

[3] O. Balci. *Verification, validation and accreditation of simulation models*, In: Andradóttir, S.; Healy, K.J.; Withers, D.H.; Nelson, B.L. (eds.); Proc. 1997 Winter Simulation Conf., IEEE Computer Society, Washington DC, USA, 1997, pp. 135-147.

[4] O. Balci. *Quality Assessment, Verification and Validation of Modeling and Simulation Applications*, In: R.G. Ingailis, M.D. Rossetti, J.S. Smith, B.A. Peters, eds., Proc. 2004 Winter Simulation Conf., USA, 2004, pp. 122-129.

[5] T.I. Ören. *Concepts and criteria to assess acceptability of simulation studies: a frame of reference*. In: Adam, N.: *Simulation Modeling and Statistical Computing*, (guest editor), Communications of the ACM, Vol. 24, No. 4, 1981, pp. 180-188

[6] D. Brade. *Enhancing modeling and simulation accreditation by structuring verification and validation results.* In: Joines, J.A.; Barton, R.R.; Kang, K.; Fishwick, P.A. (eds.): Proc. 2000 Winter Simulation Conf., Soc. for Modelling and Computer Simulation International, La Jolla CA, USA 2000, pp. 840-848.

[7] D. Brade. *A Generalized Process for the Verification and Validation of Models and Simulation Results.* Dissertation submitted for award of the degree Dr. rer. nat., Fakultät für Informatik, Universität der Bundeswehr München, Germany, 2003.

[8] T. Ericsen. *Power electronic building blocks – a systematic approach to power electronics.* In: Proc. IEEE Power Engin. Summer Meeting, Vol. 2, 2000, pp. 1216-1218.

[9] T. S. Ericsen. *Physics based design, the future of modeling and simulation*, Acta Polytechnica, Vol. 45, No. 4, 2005, pp. 59-64.

[10] W.P.J. Visser, M.J. Broomhead. *GSP: a generic object-oriented gas turbine simulation environment*, National Aerospace Laboratory, NLR, Flight Division, Technical Report NLR-TP-2000-267, Amsterdam, The Netherlands, 2000.

[11] J.H. Bigelow, P.K. Davis. *Implications for Model Validation of Multiresolution Multiperspective Modeling (MRMPM) and Exploratory Analysis*, RAND Monograph Report, RAND, Santa Monica CA, USA, 2003.

[12] M.I. Smith, D.J. Murray-Smith, D. Hickman. *Mathematical and computer modeling of electro-optic systems using a generic modeling approach*, Journal of Defense Modeling and Simulation, Vol. 4, No. 1, 2007, pp. 3-16.

[13] M.I. Smith, D.J. Murray-Smith, D. Hickman. *Verification and validation issues in a generic model of electro-optic sensor systems*, Journal of Defense Modeling and Simulation, Vol. 4, No. 1, 2007, pp.17-17.

[14] U.K. Office of Science and Technology, *Technology Foresight: Progress Through Partnership*, *Volume 12 (Defence and Aerospace)*, HMSO, 1995.

[15] Workshop on Foundations for Verification and Validation in the 21st Century, October 22-23, 2002, John Hopkins University, Laurel MD, 2002.

[16] Foundations ’04 Workshop for Verification, Validation and Accreditation in the 21st Century, October 13 – 15, 2004, Arizona State Univ., Tempe, Arizona, USA, 2004.

[17] E.H. Page, W.H. Lanceford. *Architectural principles for the US Army Simulation and Modeling for Acquisition Requirements and Training (SMART) Initiative.* In: B.A. Peters, J.S. Smith, D.J. Madeiros, M.W. Rohren, eds., Proc. 2001 Winter Simulation Conf., IEEE Computer Society, Washington DC, USA, 2001, pp. 767-770.

[18] T. Ericsen. 2005 IEEE Vehicle Power and Propulsion Conference, September 7-9, 2005, Chicago, Illinois, USA, Keynote presentation, Abstract. (http://vpp.iit.edu/VPPC2005ProgramBook_Distribution.pdf).

[19] M.B. Tischler. *System identification requirements for high bandwidth rotorcraft flight control systems*, J. Guidance, Control and Dynamics, Vol. 13, No. 5, 1990, pp 835-841.

[20] M.B. Tischler. *System identification methods for aircraft flight control development and validation*. In: M.B. Tischler, eds., Advances in Aircraft Flight Control, Taylor and Francis, London, UK, 1996, pp 35-69.

[21] D.J. Murray-Smith. *Methods for the external validation of continuous system simulation models: a review*, Mathematical and Computer Modelling of Dynamical Systems, Vol. 4, No. 1, 1998. pp 5-31.

[22] SCS Technical Committee on Model Credibility, *Terminology for model credibility*, Simulation, Vol. 32, No. 3, 1979, pp 103-104.

[23] D.J. Murray-Smith. *The validation of computer-based models in engineering: some lessons from computing science*, Acta Polytechnica, Vol. 41, No. 4-5, 2001.

[24] R.G. Sargent. *Verification and validation of simulation models*. In: F.E. Cellier, ed., Progress in Modelling and Simulation, Academic Press, UK,1982, pp 159-169.

[25] W.L. Oberkampf, T.G. Trucano, C. Hirsch. *Verification, validation and predictive capability in computational engineering and physics*, Invited Paper, Workshop on Foundations for Verification and Validation in the 21st Century, October 22-23, 2002, John Hopkins University, Laurel MD, 2002.

[26] F.M. Hemez. *The myth of science-based predictive modelling*, Foundations `04 Workshop for Verification, Validation and Accreditation (VV&A) in the 21st Century, October 13-15, 2004.

[27] O. Balci, R.E. Nance, J.D. Arthur, W.F. Ormsby. *Expanding our horizons in verification, validation, and accreditation research and practice.* In E. Yücesan, C.-H. Chen, J.L. Snowdon, J.M. Charnes, eds., Proc. 2002 Winter Simulation Conference, pp. 653-663.

[28] D.J. Cloud, L.B. Rainey, eds., *Applied Modeling and Simulation: An Integrated Approach to Development and Operation.* McGraw-Hill, New York, USA, 1998.

[29] D. Brade, A. Köster. *Risk-based validation and verification levels definition*, Proc. European Simulation Interoperability Workshop, London, UK, 2001, Simulation Interoperability Standardization Organization.

[30] D. Brade, R. Maguire, H.-B. Lotz. *Arguments-based credibility levels*. In: Proc. SISO Spring Simulation Interoperability Workshop, Orlando FL, USA, 2002.

[31] B.P. Zeigler, H. Praehofer, T.G. Kim. *Theory of Modeling and Simulation*, Second Edition, Academic Press, San Diego CA, USA, 2000, pp 367-389.

[32] E.R. Tufte. *Visual Display of Quantitative Information*, Graphics Press, Cheshire CT, USA, 1983.

[33] A.G.N. Grant. *Data to Polygon Transformation and Comparison: A Technique for Model Validation and Related Applications*. PhD Thesis, University of Glasgow, UK, 2006.

[34] J.J. Zenor, D.J. Murray-Smith, E.W. McGookin, R.E. Crosbie. *Development of a multi-rate simulation model of an unmanned underwater vehicle for real-time applications.* In: I. Troch, F. Breitenecker, eds., Proc. MATHMOD’09 Vienna, CD Volume, ARGESIM, Vienna, Austria, 2009, pp. 1950-1957.

[35] R.J. Malak Jr., C.J.J. Paredis. *Foundations of validating reusable behavioural models in engineering design problems.* In: R.G. Ingalls, M.D. Rossetti, J.S. Smith, B.A. Peters, eds., Proc. 2004 Winter Simulation Conference, USA, 2004, pp. 420-428.

[36] G.T. Mackulak, F.P. Lawrence, T. Colvin. *Effective simulation model reuse: a case study for AMHS modelling*. In: D.J. Medeiros, E.F. Watson, J.S. Carson, M.S. Manivannan, eds., Proc. 1998 Winter Simulation Conf., IEEE, Piscataway, NJ, USA, 1998, pp. 979-984.

[37] R.E. Skelton, ed., *Model Validation for Control System Design*, Special Issue, Mathematical Modelling of Systems, Vol. 3, No. 1, 1997.

[38] L. Ljung, L. Guo. *Classical model validation for control design purposes*, Mathematical Modelling of Systems, Vol. 3, No. 1, 1997, pp 27-42.

[39] Smith, R.S.; Doyle, J.C.: *Model validation - a connection between robust control and identification*, IEEE Trans. on Automatic Control, Vol. 37, No. 7, 1992, pp 942-945.

[40] Ljung, L.: *System Identification: Theory for the User*, Prentice Hall, Englewood Cliffs NJ, USA, 1987.

[41] D.J. Murray-Smith. *Issues of model accuracy and external validation for control system design*, Acta Polytechnica, 40, 3, 2000, pp 8-13.

[42] Defense Modeling and Simulation Office, *Verification, Validation and Accreditation (VV&A) Recommended Practices Guide (RPG)*, Alexandria VA, USA. 2000.

[43] H. Cross. *Engineers and Ivory Towers* In: R.C. Goodpasture, ed., McGraw-Hill, New York, USA, 1952.

[44] J. Kocijan, A. Girard, B. Banko, R. Murray-Smith. *Dynamic systems identification with Gaussian Processes*, Mathematical and Computer Modelling of Dynamical Systems, Vol. 11, No. 4, 2005, pp 411-424.

Corresponding author: David Murray-Smith
Dept. of Electronics and Electrical Engineering
University of Glasgow
Glasgow G12 8LT, Scotland, United Kingdom
djms@elec.gla.ac.uk