

Beyond VVA: Model and simulation users

John Barlow

Australian Catholic University
john.barlow@acu.edu.au

Abstract: Much is quite rightly written about the importance and significance of robust VVA (Verification, Validation and Accreditation) systems as part of the modelling and simulation development process. While these processes are clearly significant and widely required in establishing the overall fidelity, veracity and, in many instances, supporting the registration for use of a model or simulation, there is also another significant matter that needs to be examined. That is, the need or otherwise for users of these products to be either suitably prepared and trained, or at least provided with suitable information to indicate the type of product they seek to interact with. This paper will briefly outline VVA, followed by a critical examination of the importance of user training in the overall successful implementation of model and simulation products with particular reference to an examination of ethical and professional responsibilities of model and simulation developers, as well as the users of these products.

Keywords: accreditation, ethics, persuasion, validation, verification

1. INTRODUCTION

There is no doubt that computer simulations are used extensively by a wide range of individuals to assist them in such activities as their day to day business and professional decision making in areas such as economic and financial forecasting, environment and ecology management, and water resource management and monitoring, as well as assessing and reducing risk, and staff training. The significance of the verification, validation and accreditation (VVA) of such computer simulation tools cannot be underestimated. Clearly the rigour and veracity of any simulation tool is very much dependent upon the systems and procedures that have been put in place to establish its credibility. In this context there is no doubt that numerous well practised, accepted and proven procedures are commonly utilised to establish such an outcome. However the development and establishment of the credentials and credibility of any simulation tool is only the first part of the process of such a tool being used constructively within its limitations and for its intended purpose. Importantly the second part of the process is establishing the competence of the various users of these simulation tools.

Clearly one of the most important reasons modelling and simulation products are used is to assist and support confident decision making by their users. This can play out across a number of levels ranging for example from deciding what is the next most astute decision to make to win the next point while playing a simulated game for recreational interests to far more serious instances involving decision making based on complex prototyping or financial modelling where the consequences of poor decision making may impact significantly on the interests of the user and many other individuals both directly and indirectly. While the successful recreational use and enjoyment of games which utilise simulations may be enhanced by a fundamental level of expertise and skill, it does not necessarily involve or indeed require appropriate and accredited training. However the use of increasingly sophisticated simulation products more generally may require appropriate training, knowledge and an appreciation of the intended capacity and limitations of these products. This is an important consideration given that the output of modelling and simulation products can only at best be interpreted as conditional prediction, the confidence of which is underpinned by numerous factors which may or may not be known by their users.

While recognising the importance and significance of VVA, this paper seeks to examine issues associated with the characteristics and intent of computer simulation product users. In particular it will canvass those issues which may significantly impact on and influence the reliable and valid use of such products. The concept of modelling and simulation product VVA will be briefly examined followed by related issues of computers and simulation products as persuasion tools in the wider community. In addition a suggested classification of the nature of computer simulation product users, including both professional and non-professional users, considered important for this discussion will be briefly outlined. This will be followed by a discussion of modelling and simulation product user issues identified as relevant to the appropriate use and interpretation of such products. In doing so the intention of this paper is simply concerned with flagging the ever present potential for the inappropriate use or misuse of what are otherwise valid and accredited simulation products.

2 VERIFICATION, VALIDATION AND ACCREDITATION

While the exact nature of models and simulations can be contentious (Eason, Rosenberger, Kokalis, Selinger and Grim, 2007), in this paper a model is taken to be a representation or abstraction of an existing situation, product, or idea while a simulation is taken to be the manipulation and investigation of a model's behaviour using a range of real or abstract data. Indeed Winsberg (2006, p 4) observes that simulations "are often constructed precisely because data about the systems they are designed to study are sparse." While the use of the terms verification and validation are often used interchangeably in some of the modelling and simulation literature, the fundamental purpose of verification and validation is to ensure that a model and simulation are accurate. Put simply verification is the process of ensuring that the conceptual idea of a model has been accurately transformed into a computer model while validation is the process of ensuring that the model is sufficiently accurate for its specified purpose.

Essentially validation is primarily a procedural process which needs to be constantly addressed and monitored during model and simulation development and evaluation to ensure that the final product which is implemented for use is as fundamentally rigorous as possible. Validation procedures and tests need to be seen as quality assurance measures which are ideally put in place to determine the veracity and rigour of all the components which when joined together form the final model and simulation product. Clearly it needs to be acknowledged that a number of issues may impact on establishing the credentials of any model and simulation product. These include but are not necessarily limited to issues such as establishing the context

and veracity of the situation, product or idea the model is designed to represent, the quality of the data available for use during both the development and use of the product, perceptions of modelling and simulation developers and users, as well as the time available for testing and evaluating the performance of and confidence in the output provided by the final product.

Issues associated with VVA have been comprehensively examined in the literature. A number of exemplars are provided to illustrate complex issues which can impact on an understanding of VVA. This is not meant to be an exhaustive literature review but merely examples of the kinds of issues which have been raised in recent years. Naylor and Finger (1967, p B-93) make the point that in order to verify or validate a model, that is prove that it is true, two considerations need to be put in place.

1. A set of criteria needs to have been established for distinguishing between models which are ‘true’ and those models which are ‘not true’, and
2. There is the ability to apply the criteria to the models.

Further they also note Popper’s (1959) suggestion that since it may be difficult to completely agree on a set of suitable criteria for testing a model’s validity, it may be more appropriate to concentrate on establishing the degree of conformation of a model rather than whether or not the model has been verified. In this context Sterman (2000, p 845) has argued that emphasis should be placed on establishing the confidence with which a model can be used for its proposed purpose rather than whether or not it has been verified and validated. More broadly Kleijnen (2000) observed that VVA has many facets including both philosophical and mathematical-statistical problems. In essence the fundamental philosophical issue relate to addressing the question “What is the truth?” a model and simulation claim to represent, and how can this be recognized while statistical issues relate to identifying, using and critically interpreting the results of appropriate statistical tools during and after the validation process.

Robinson (2004, p 209) also notes for example that validation in particular operates at a number of interrelated levels. These are more broadly set out in table 1.

Forms of validation	Description
<i>Conceptual model validation</i>	Concerned with determining that the content, assumptions and simplifications of the proposed model are sufficiently accurate for the purpose at hand.
<i>Data validation</i>	Concerned with determining that the contextual data and the data required for model realisation and validation are sufficiently accurate for the purpose at hand.
<i>White-box validation</i>	Concerned with determining that the constituent parts of the computer model represent the corresponding real elements with sufficient accuracy for the purpose at hand (micro analysis).
<i>Black-box validation</i>	Concerned with determining the overall model represents the real world with sufficient accuracy for the purpose at hand (macro analysis).
<i>Experimental validation</i>	Concerned with determining that the experimental procedures adopted are providing results that are sufficiently accurate for the purpose at hand.
<i>Solution validation</i>	Concerned with determining that the results obtained from the model of the proposed solution are sufficiently accurate for the purpose at hand.

Table 1: Forms of validation (Robinson, 2004)

As an outcome of Foundations ’02, Pace (2004, p167) outlined two sets of challenges facing the modelling and simulation community in the context of the validation process: management challengers and research challengers.

Management challenges were identified as the:

1. use of qualitative assessment procedures undertaken by individuals who may not necessarily possess appropriate credentials and an appreciation of formal methods,
2. development of appropriate and effective ‘lightweight’ variants of more complex formal assessment procedures, and
3. development of appropriate cost effective resource information.

Research challenges were identified as the need:

1. for scientifically rigorous methods for making predictions (inference) about simulation results in the application domain where data do not exist,
2. to establish the rigour of emerging adaptive programming techniques including artificial intelligence, expert systems, and genetic algorithms,
3. to develop methods that can minimise the potential misuse of complex distributed computer model and simulation products, and
4. to appreciate that where simulations involve humans, interactions among simulation characteristics, the people involved and appropriate simulation users require further research..

Sargent (2007, p 125) has also provided a summary of four basic approaches he suggests can be adopted to establish a model and simulation product's validity. These are:

1. The model development team makes the decision as to whether a simulation model is valid,
2. User(s) of the model are involved with the model development team in deciding whether a model is valid,
3. A third party is used to undertake a 'independent verification and validation' (IV&V) to decide whether a model is valid,
4. Scoring machines, where scores assigned to various aspects of the model and simulation development validation processes are determined subjectively and combined to determine an overall score for the simulation model.

More recently Balci (2010), among others, has noted that verification is concerned with transformational accuracy while validation is concerned with behavioural or representational accuracy.

Finally the accreditation of a modelling and simulation product generally involves an organisation or person undertaking a rigorous review and assessment of whether the product is sound and can be used for the purpose it was specifically developed for. The accreditation process generally incorporates a range of tests which are designed to assess the veracity of the product across a range of factors including both the model and simulation product itself, as well as the input data and output provided by the product. The fundamental reason accreditation is such an important component of the overall modelling and simulation product development process is that a successful accreditation outcome can obviously enhance the products credibility within the context and purpose it has been designed to function in.

3 BEYOND VVA

Clearly VVA is generally accepted as a significant part of the modelling and simulation development process. However it is important to appreciate that the significance and impact of computer models and simulation products may extend well beyond their immediate and intended audience. In this context it is equally important to appreciate the nature of computer simulations as sophisticated persuasion tools.

Consider this phrase: 'According to our computer modelling ...' often implies either implicitly or explicitly what is to follow has some credibility and trustworthiness, and for that matter even authority and therefore as a consequence should be an important and influential consideration when weighing up all the factors associated with a particular situation and then making consequential decisions. In this context weather forecasting, and to a limited extend financial planning are probably now two of the most common examples where the wider community is commonly exposed to the output of this 'credible' phrase, resulting in individuals making decisions about how they will conduct their short and long term affairs which may or may not result in significant consequences. To be fair though in the context of weather forecasting it would be reasonable to assert that while realising the accuracy of such forecasts is increasingly improving, many members of the community may still choose to give weather forecasters the benefit of the doubt!

Fogg, Cuellar and Danielson (2007, p 134) suggest the term *persuasion* can be defined as a noncoercive attempt to change attitudes and behaviours. Further they assert that this definition has three important features in that persuasion is noncoercive, that is it does not involve force, it requires an attempt to change involving intentionality, and it deals with attitude or behaviour change which they contend is fundamental to the study of computers as persuasive technologies. In this context they identified three types of computer simulations as being relevant to persuasive technologies:

1. Simulated cause-and-effect scenarios
2. Simulated environments
3. Simulated objects

Clearly the three types of computer simulations identified above reflect a wide range of products produced and used by the modelling and simulation community. The important consideration here is computer model and simulation products are now more than ever directly and indirectly accessible to increasing numbers of the broader community and quite rightly provide them with a wide range of opportunities and support systems which they commonly use to directly and indirectly influence and support their private, professional and business decision making.

However not all simulation users necessarily possess the technical expertise or knowledge required to understand the issues involved in the development and responsible use of a modelling and simulation product. As a consequence they may not always be aware of the limitations or accredited use of such products. Broadly users of modelling and simulation products may be classified in a number of ways. For example:

1. *Expert v non-expert users.* Expert users possess appropriate and relevant knowledge and expertise associated with the limitations and capabilities of a simulation product while non-expert users may not.
2. *Direct v indirect users.* Direct users provide input, manipulate and interpret the output of a simulation product while indirect users may rely on direct users to provide them with such information.

Clearly expert users may also be direct users while non-expert users may also be indirect users. Moreover the context of use will also determine the status of the user. For example an individual may be considered an expert simulation user in one context but a non-expert in another where their domain knowledge is either limited or even non-existent.

For the purpose of this paper however users will be described under three levels.

1. *First level users* – the ‘professionals’ who are trained to use the simulations and can ‘accurately’ interpret a simulation output and what may be the implications of the output. This level of users includes clients and other persons who have relevant expertise and more importantly have had access to relevant training in the use of the simulation tool.
2. *Second level users* – untrained users who ‘shop’ for an appropriate simulation tool that will assist their decision making. This level of users can include a range of people including for example those individuals who may use a superannuation calculator to decide when they might retire.
3. *Third level users* – users who use information supplied by a third party to make decisions. This level of users may include ‘passive’ recipients of simulation outputs including, for example, individuals who use weather forecasts provided by commercial media personnel who are not necessarily trained and qualified meteorologists.

First level users are those individuals who would be generally seen and recognised as the classical professional users of modelling and simulation products. They are respected for their professional expertise, knowledge, training and perhaps most importantly their authority within the context of the model and simulation product’s domain and wider expert community. Ideally these should be the users who have some knowledge of the issues associated with the development and implementation of the modelling and simulation products they utilise in the exercise of their professional and employment responsibilities.

By way of contrast second and third level users are not necessarily professional and more than likely not necessarily prepared or even interested in undertaking appropriate training in the use and interpretation of the modelling and simulation product output. Moreover they may not be aware or see the need to be aware of the specifications and limitations underpinning the development of simulation products they may choose to utilise. This is not to suggest they are necessarily unprofessional or irresponsible individuals who always deliberately and irresponsibly use modelling and simulation products in inappropriate and unprofessional circumstances. Indeed they may be operating in misguided good faith. More importantly however Prazeres, Clarke, Hand and Kim (2007) assert that given that simulation tools generate large amounts of complex data, the translation of this data useful for decision making is problematic. Clearly then the experience and knowledge of simulation users is a critical factor in their ability to utilise a simulation tool. This involves a range of issues including their knowledge of what kind of data should be used to populate a simulation as well as their capacity to astutely analyse and interpret the output obtained from a simulation package. In addition, consideration needs to be taken account of their level of training, access to appropriate technical support, the accountability of the consequences of their decision making, as well as taking responsibility for the decisions made on the basis of a simulation output. That is the simulation is not necessarily the problem, the inappropriate use of the simulation is the problem. Bahill and Henderson (2005) have provided a

comprehensive discussion highlighting instances where requirements development, verification and validation 'errors' by professional personnel have contributed to unfortunate outcomes. For example the satellite thrusters for the Mars Climate Orbiter were manufactured using imperial units while the satellite operator modelled the system using metric units resulting in the orbiter entering the Mars atmosphere and burning up. Even experts make mistakes! More generally though Lowell, Christy, Pelizaro, Day, Barlow, O'Leary, and Pettit, (2009) have outlined issues associated with the development and use of a landscape management spatial viewer which can be accessed by non expert users. They noted the need to be aware of the different vocabularies, backgrounds and capabilities of the parties associated with the development and use of the viewer as well as the need to be aware of exactly what the output of such a tool is intended to provide. In this instance, the spatial viewer is intended to provide relative rather than absolute indications of actual and potential impacts of land use change. Clearly inexperienced, novice or inappropriate users would not necessarily be aware of this.

The issue then is how can second and third level users be managed and supported? Within what might be described as the professional community of model and simulation developers and users a number of regulatory mechanisms are commonly already in place including licensing arrangements which restrict and control access to and use of many modelling and simulation products. These licensing arrangements can also include access to appropriate training provided by accredited training personnel and regular updates of the product as these become available. In the broader community of users such mechanisms may also be in place but do not always deliver a rigorous management regime which may or may not even be enforceable.

One way forward in this instance may be the adoption or at least the consideration of trialling a scheme similar to that used throughout the food nutrition industry: the traffic light monitoring system (Borgmeier and Westenhoefer, 2009). In brief, green implies safe use amber implies cautious use and red dangerous use. If such a system was adopted and assigned to the myriad of modelling and simulation products which are more generally available to the wider community it may contribute to the development of a mechanism for self regulation and informed use of such products. For example, superannuation calculators which are commonly provided on superannuation fund websites are currently quite often provided with some form of disclaimer about their use and the information they generate which is obviously very appropriate and perhaps even required by law. However an amber traffic light might also be provided which could indicate that the use of such a product should be undertaken with care. The very real advantage of considering a traffic light system in this instance is that the required training is minimal. Such a system would not guarantee that modelling and simulation products were not always used inappropriately, but it may certainly provide the potential for a generally universal communication classification structure which most if not all individuals are aware of, and as a consequence support the development of both an informal and formal self regulated management regime.

Finally having provided a suggestion here is only in part the first step in providing a solution. Clearly mechanisms would need to be put in place for determining which colour traffic light and in what circumstances it should be assigned to individual modelling and simulation products. No doubt this can be addressed if a broad consensus could be established to manage the inappropriate use of modelling and simulation products.

4 CONCLUSION

This paper has provided a brief outline of some of the issues associated with the verification, validation and accreditation of modelling and simulation products with a few to examining issues associated with the inappropriate use of such products. The notion of computer simulations as persuasion tools was canvassed as was the characteristics of modelling and simulation product users. It was suggested in instances where modelling and simulation products are potentially exposed to inappropriate use, a traffic light monitoring system might be developed which may assist the development of a more self regulated management approach to such problems.

5 REFERENCES

- Balci, O. (2010). *Golden Rules for Verification, Validation, Testing, and Certification of Modeling and Simulation Applications*, SCS M&S Magazine, No. 4.
- Bahill, A. T. And Henderson, S. J. (2005). *Requirements Development, Verification, and Validation Exhibited in Famous Failures*, Systems Engineering, Vol. 8, No. 1, pp 1 – 14.
- Borgmeier, I. and Westenhoefer, J. (2009). *Impact of different food label formats on healthiness evaluation food choice of consumers: a randomized-controlled study*, BMC Public Health, Vol. 9: 184.
- Coon, J. E., Kusch, M. G., Rowland, M. C. and Cunningham, J. R. (1998). *Simulation Software and Engineering Expertise: A Marriage of Necessity*, AIChE Meeting, <http://www>
- Cockroft, J, Ghauri, S., Samuel, A. And Tuohy, P. (2009). *Complex energy simulation using simplified user interaction mechanisms*, 11th International IBPSA Conference, pp 1962-1967.
- Eason, R., Rosenberger, R., Kokalis, T., Selinger, E. And Grim, P. (2007). *What kind of science is simulation?*, Journal of Experimental & Theoretical Artificial Intelligence, Vol. 19, No. 1, March, pp 19 – 28.
- Fogg, B. J., Cuellar, G. and Danielson, D. (2007). Motivating, influencing and persuading users, <http://bjfogg.com/hci.pdf>, accessed 18 June 2011.
- Henderson, S. G., Biller, B., Hsieh, M. –H., Shortle, J, Tew, J. D. and Barton, R. R. (eds), (2007). *Proceedings of the 2007 Winter Simulation Conference*.
- Johansson, B., Jain, S., Montoya-Torres, J., Hugon, J., and Yücesan (eds) (2010). *Proceedings of the 2010 Winter Simulation Conference*.
- Kleijnen, J. P. C. (2000). *Strategic Directions in Verification, Validation, and Accreditation Research: a Personal View*, panel session WSC 2000.
- Lowell, K., Christy, B., Pelizaro, C., Day, G., Barlow, K., O’Leary, G. and Pettit, C. (2009). *Making results of complex systems-based landscape models more accessible to non-expert users*, 18th World IMACS/MODSIM Congress, <http://mssanz.org.au/modsim09>, accessed 18 June 2011.
- Naylor, T. H. and Finger, J. M. (1967). Verification of Computer Simulation Models, *Management Science*, Vol. 14, No. 2, October, pp B-92 – B-101.
- Pace, D. K. (2004). Modeling and Simulation Verification and Validation Challengers, Johns Hopkins APL Technical Digest, Vol. 25 No. 2, pp 163 – 172.
- Popper, K. (1959). *The Logic of Scientific Discovery*, Basic Books, New York.
- Prazeres, L., Clarker, J. A., Hand, J. and Kim, J. (2007). Delivering Building Simulation Information via New Communication Media, *Proceedings: Building Simulation 2007*, pp 1499 – 1505.
- Robinson, S. (2004). *Simulation The Practice of Model Development and Use*. John Wiley, Chichester.
- Sargent, R. G. (2007). *Verification and validation of simulation models*, in Henderson, S. G., Biller, B., Hsieh, M. –H., Shortle, J, Tew, J. D. and Barton, R. R. (eds), pp124 – 137.
- Sargent, R. G. (2010). *Verification and Validation of Simulation Models*, in Johansson, B., Jain, S., Montoya-Torres, J., Hugon, J., and Yücesan (eds), pp 166-183.
- Sterman, J. D. (2000). *Business Dynamics: Systems Thinking and Modeling for a Complex World*, McGraw-Hill.
- Thomson, A. J. and Schmoltdt, D. L. (2001). *Ethics in computer software design and development, Computers and Electronics in Agriculture*, Vol. 30, pp 85 – 102.
- Winsberg, E. (2006). *Models of success versus the success of models: reliability without truth*, Synthese, No. 152, pp 1 – 19.