Model Output Uncertainty in Risk Assessment

Enrico Zio, Terje Aven

Model Output Uncertainty in Risk Assessment

TERJE AVEN[*1] and ENRICO ZIO[2]

[1]University of Stavanger, Norway

[2]*Chair on Systems Science and the Energetic Challenge, European Foundation for New Energy - Electricité de France, at ÉcoleCentrale Paris - Supelec, France and Politecnico di Milano, Italy*

(Received on June 01, 2012 and revised on May 28, 2013

Abstract: In the context of risk assessment, we focus on the prediction of an unknown quantity Z whose value is realised in the future, and for which experimental data are not available. We deal with the issue of the uncertainty associated to the difference between the output of the model used for the prediction of Z and the true unknown value of Z itself. Accepted principles and methods for handling this uncertainty in the specific conditions of risk assessment are still lacking. Through the paper we seek to contribute by:

i) making a clear distinction between model output uncertainty (epistemic uncertainty about the differences between the true values of the output quantities and the values predicted by the model) and sources of model output uncertainty, from incomplete/imprecise knowledge on the values of the parameters of the model to model assumptions, simplifications and approximations introduced in the model,

ii) distinguishing between model output uncertainty, structural model uncertainty and parameter (model input quantities) uncertainty,

iii) establishing explicit links between the different purposes of modelling and risk assessment, discussing how model output uncertainty should be treated in the different instances.

We argue that in risk assessment, quantification of model output uncertainty serves for the qualification and acceptance of the models used, whose outputs feed the following risk-informed decision making process.

Keywords: *Risk assessment, model uncertainty, model error; structural model uncertainty, parameter uncertainty*

1. Introduction

Risk assessment is based on models of the phenomena, events and systems of interest (the “real world”).Model uncertainty is, then, a relevant concept, which often arises in the consideration of the quality and reliability of the assessment, but with different meanings. It is common to consider model uncertainty when there are alternative plausible hypotheses for describing the phenomena, events and systems of interest [30,35].But model uncertainty is also addressed for a single model, which is generally accepted but not completely validated, or it is conceptually accepted and validated but its implementation is of uncertain quality, or it is recognized to only partially cover the relevant aspects of the problem, or it is composed of submodels of different degrees of accuracy and credibility [10]. Finally, model uncertainty is also analyzed to describe the differences between the actual, real world output values and the output values of the model [16,23,28].

The purpose of the present paper is to contribute to framing and better understanding the problem of model uncertainty in the context of risk assessment by a fundamental rethinking of the concept itself. Whereas much of the existing views have been founded on the probability of frequency approach to risk assessment (implying the existence of probabilistic models of the aleatory uncertainties, with unknown parameters subject to epistemic uncertainties [11]), our view expands onto a broader framework, to capture both the probability of frequentist set-up and the modelling supported by subjective probability assignments (or alternative approaches for expressing epistemic uncertainties, like imprecise probabilities).

Secondly, we specify the concept in relation to the objectives of the modelling and risk assessment, discussing how uncertainty should be treated in the different instances. For the objectives, four categories are often distinguished in industrial practice [8]:

- To reach a required quality level of a model by validation, for its certified use (accredit)
- To understand the influence of uncertainties on the results of the analysis and rank their importance so as to guide additional efforts (measurements, research, etc.) for uncertainty reduction (understand)
- To compare performances of alternative system designs, operation modes and maintenance policies, for “optimal” choices (select)
- To demonstrate compliance of a system, process, procedure with regulatory criteria (comply).

Our focus is on the evaluation of the model output uncertainty, i.e. the uncertainty associated to the difference between the model output values and the true values of the quantities of interest for the risk assessment, for example the number of fatalities in relation to accidents in an hazardous process plant or the number of occurrences of a major accident on an offshore installation leading to a major oil spill. The situations we are interested in are those characterised by a lack of experimental data for comparing the “goodness” of and validating the models used for evaluating these quantities. The quality of the models, then, needs to be based on a good understanding and proper argumentation with respect to the events and phenomena of interest. Where available, detailed models are used to describe specific features of the phenomena under considerations, and these can often be checked and validated through experimental data. However, this type of models where experimental data are available is outside the scope of the present paper. The present work addresses the problem at the system level, where the quantities of interest are defined, and where experimental data are not available.

2. A Formal Set-up for Defining and Understanding Model Uncertainty

We consider a problem in which by way of a model we seek to know a quantity Z, whose value is realised in the future. It could for example represent the failure time of a system, or the volume of oil that reaches vulnerable areas in case of oil spill. A model G(X) is developed for this purpose, whose output depends on input quantities (parameters) X and the function G.

Let $\Delta_G(X)$ be the differencebetween the model prediction, G(X), and the true future value Z, *i.e.*, $\Delta_G(X)= G(X) - Z$. We refer to $\Delta_G(X)$ as the model error and to the uncertainty on its value as model output uncertainty.

Note that according to this definition, model output uncertainty is not the same as the model error $\Delta_G(X)$: it is actually the epistemic uncertainty of it, and may be assessed by some measures of epistemic uncertainty, for example subjective or imprecise probabilities. The model output uncertainty and its measurement are considered in relation to the magnitude of the model error. For example, if we use probability to represent uncertainty, this is expressed by the probability distribution of the error. Hence if we have a situation in which the probability mass is concentrated around a specific value of the error, the description of the model output uncertainty would present a peak in the distribution in correspondence of such error, and an appropriate judgment can be made onitssignificance for the purposes of the risk assessment and associated decision making.

Model output uncertainty results from the combination of two components: structural model uncertainty and parameter (input quantity) uncertainty, defined as follows:

Structural Model Uncertainty: Uncertainty about the difference$\Delta_G(X_{true})$, when the true value $X_{true}$of the input quantity X is known. In other words, the structural model uncertainty expresses the uncertainty (lack of knowledge) on the model output error when we can ignore the uncertainties about the parameters (input quantities) X (hence the uncertainties are due to the structure G of the model, alone).

Input Quantity (parameter) Uncertainty: Uncertainty (due to lack of knowledge) about the true value of the input quantities X.

Hence model output uncertainty results from both structural model uncertainty and parameter uncertainty. Note that various authors call model uncertainty the part of structural model uncertainty, keeping separate the parameter uncertainty.

As for the model output uncertainty, the structural model uncertainty and the parameter uncertainty (with their measurements) are considered in relation to the magnitude of the relevant unknown quantities, $\Delta_G(X_{true})$ and X, respectively.

Sources of structural model uncertainty stem from actual "gaps" in knowledge which can take the form of poor understanding of phenomena that are known to occur in the system, as well as complete ignorance of other phenomena. This type of uncertainty can lead to "erroneous" assumptions regarding the model structure.Further uncertainty in the quantitative mathematical models derives from approximations and simplifications introduced in order to translate the conceptual models into tractable mathematical expressions. The implementation of the mathematical models into computational codes introduces additional uncertainty caused by possible coding errors, discretization schemes, computational limitations, etc. Various methods have been developed for the identification of this uncertainty (see e.g. Oberkampf et al. [27]). Notice that it is possible to somewhat reduce this uncertainty by means of good Quality Assurance practice, e.g. by validation and sufficient computing power to render discretization errors negligible.

We argue that our view on the uncertainty of the model error is the one of interest in the practice of quantitative risk assessment, for analysing the performance of the model G(X) which we use for predicting the quantity Z of interest for the risk assessment: this requires accounting for both the features of the model structure and the variability of the input quantity values, and the related lacks of knowledge. Assessing $\Delta_G(X_{true})$ when conditioning on the true value of X, is also of interest, but what counts in the end is all we know, and do not know about the error made when using the output of the model G(X) in place of Z.

If we use probabilities to express the model output uncertainty, we have by the law of total probability:

$$P(\Delta_G(X) \leq z) = \int P(G(x)-Z \leq z \mid X=x)\, dH(x),$$

where, H is the probability distribution of X. Hence, the model output uncertainty is the structural model uncertainty weighted over the distribution of the input quantities.

2.1 Examples

Example 1: The Poisson Model for the Occurrence of Some Type of Events

Consider the Poisson model for describing the occurrence of an event in the future. Let $p(n|\lambda,t)$ be the probability distribution of the random variable N_t representing the number of times the event occurs in the interval [0,t], where λ is the event occurrence rate, *i.e.*,

$$P(N_t =n \mid \lambda, t) = p(n|\lambda, t) = (\lambda t)^n e^{-\lambda t}/n!, \quad n=0,1,2,\ldots .$$

We denote $p_0(n|t)$ the true distribution of N_t and seek to verify that its 95^{th}quantile, n_{95}, complies with a given regulatory threshold n_M, i.e. $n_{95} \leq n_M$.

Using the (Z,X,G) terminology introduced above, $Z = n_{95}$, $X = \lambda$ and $G(\lambda)$ is the 95^{th}quantileof $p(n|\lambda,t)$.

Hence the model error $\Delta_G(X)$ is given by $G(\lambda) - n_{95}$, and the model output uncertainty is the combined structural and parameter uncertainty (due to lack of knowledge) about this error. The structural model uncertainty is uncertainty about $G(\lambda_0) - n_{95}$when we know the true value λ_0 of the parameter λ. This uncertainty is a result of the poissonian representation $p(n|\lambda_0, t)$ of $p_0(n|t)$, *e.g.*,because of the assumption that the event occurrence rate is constant. The parameter uncertainty is uncertainty about the true value of λ.

Example 2: An Event Tree Model

An event tree model is developed in a risk assessment of an LNG (Liquefied Natural Gas) plant in an urban area [2], see Figure 1. The quantity of interest is Z, the number of fatalities as a result of a release.

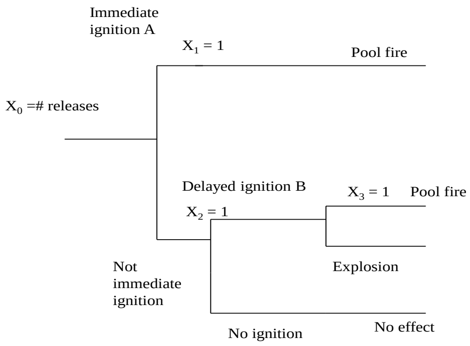

Figure 1: Event Tree for LNG Plant

Here:
X_0= number of releases (which is approximately equal to 1 if a release occurs and 0 otherwise, as we ignore the probability of two releases in the period studied)
X_1 = I(A) (I is the indicator function, which is equal to 1 if the argument is true and 0 otherwise) A: Immediate ignition
X_2 = I(B), B: Delayed ignition
X_3 = I(pool fire).
$X = (X_0, X_1, X_2, X_3)$

We see that if a release occurs, it can either result in a pool fire, an explosion or no effect, depending on the results of the branching events, immediate ignition and delayed ignition.
The model provides four scenarios:
s_1: release - A - pool fire
s_2: release - not A – B - flash (pool) fire
s_3: release - not A – B - explosion
s_4: release - not A - not B - no effect.
Assume that the number of people exposed to scenario $s_i$is v_i, where $v_1 = 0$, $v_2 = 50$ and $v_3 = 100$.
Furthermore, assume that the fraction of fatalities is d_i, where $d_2 = d_3$= 0.1.
The model G for the number of fatalities is given by
$G(X) = 5 X_0 (1- X_1) X_2 X_3 + 10 X_0 (1- X_1) X_2 (1-X_3)$,
as the number of fatalities is 5 in case of scenario 2, and this scenario occurs if $X_0(1- X_1) X_2 X_3 = 1$, and the number of fatalities is 10 in case of scenario 3, and this scenario occurs if $X_0(1- X_1) X_2 (1-X_3) = 1$.
The model error $\Delta_G(X)$ is given by the difference in the number of fatalities, G(X) - Z, and model output uncertainty is the uncertainty (due to lack of knowledge) about this error. The structural model uncertainty is uncertainty about G(X) - Z, where it is given that we know the true value of X. The parameter uncertainty is uncertainty about the true value of X: for example, the number of releases.

Example 3: A Groundwater Flow Model

Consider the groundwater flow modelling problem addressed in Zio and Apostolakis [39] and Baraldi and Zio [5], for the evaluation of the quantity of interest Z which is the hydraulic head spatial distribution (reflecting aleatory uncertainties) in the far field region of a radioactive waste repository site. Four alternative hypotheses have been considered regarding the hydro-geological properties of the media in which the repository is located and the groundwater flow mechanisms:

- Groundwater flow in porous matrix only.
- Groundwater flow in fractures only.
- Groundwater flow in matrix and fracture simultaneously.
- Groundwater flow in matrix and fracture simultaneously, but four stratigraphic layers only.

The model error $\Delta_G(X)$ is given by the differences between the distributions predicted by the models and the true spatial distribution of the hydraulic head on site, Z, and model output uncertainty is the uncertainty (due to lack of knowledge) about this error. The structural model uncertainty is uncertainty about G(X) - Z when it is given that we know the true value of X. The parameter uncertainty is uncertainty about the true value of X.

2.2 Dealing with Model Uncertainty, in view of the Objectives of the Assessment

Starting from the concept of uncertainty about the model error, we address the issue on how to deal with it, in relation to the objectives of the assessment and the causes of uncertainty. In the following we discuss this in detail using the above three examples as illustrations.

We take the understanding of accredit as the objective of reaching a required quality level of a model by validation, for its certified use. Clearly, this requires that model output uncertainty be sufficiently small, for confidence in the use of the outputs produced by the model. What is sufficiently small is of course dependent on the purpose for which the model is to be used. In practice, model accreditation stands on the evaluation of the comparison of the model predictions G(X) with the corresponding true values of Z, for establishing the level of confidence in the model predictive capability needed for the intended use of the model: the accreditation must demonstrate that in correspondence of values of X the model produces predictions of the true values of Z with the sufficient level of accuracy and the confidence required for taking related decisions. In the case of accreditation, then, the evaluation of the model output uncertainty serves the purpose of verifying the level of accuracy achieved so as to have the confidence required to make decisions informed by the outcomes of the model.

In the case that experimental data are available, there exists a wide range of statistical methods that can be used for validation in order to accredit a model. These methods include both traditional statistical analysis and Bayesian procedures [e.g. 6,15,17,21,37, 38]. However, these methods are not within the scope of the present paper, in which we consider situations with lack of data. In such cases, the validation task takes a different form, the main tool that can be used being analyst/expert argumentation based on established scientific theories and specific knowledge about the system/activity which the model being assessed intends to describe.

Model validation is often linked to model verification (and is often referred to as Verification and Validation, or simply "V&V."), which is commonly understood as the process of comparing the model with specified requirements [18,20,25,34,36. The verification part is obviously important in many contexts to produce a model that meets the specifications; however, for the purpose of the present paper we will not discuss the issue in more detail, as our analysis is concerned about the fundamental thinking that precedes the specification process.

To exemplify the situation described above within the risk assessment context of interest in this paper, consider the probability model of Example 1. To assess the model output uncertainty for purposes of accreditation of the model, amounts to justifying the model hypotheses through a set of arguments, the important ones being in this case those supporting the assumption that the probability of an event occurring in any small interval of time is the same, independent of the history of previous occurrences.

Depending on the situation, such type of arguments can be considered sufficient to conclude that the model output uncertainty is sufficiently small for the intended uses of the model. On the contrary, questions could be raised, for example, about the assumption that the occurrence rate is not dependent on time. In absence of data, it is not possible to perform a statistical analysis to study this issue any further: the risk analysts need to argue the issue and make their judgments, which may very well lead to a suggestion for an adjusted model, for example a Poisson model with time dependent occurrence rate, and a repetition of the analysis process before the related decision is taken, in a sort of sensitivity analysis “attitude”. For some uses, it may be decided that the model can be accredited even if the analyst knows that the errors could be large, as long as the analysts can justify that the model is able to reflect the key features of the phenomena studied, of interest for its use.

In the end, the results of the assessment are communicated as conditional on the model, and must be understood as dependent on the model.

Say that the analysts conclude that the assumption of a homogenous Poisson model is reasonable in the considered situation. This means that the structural model uncertainty is judged sufficiently small for the purpose of the analysis. The parameter uncertainty of the event occurrence rate λ is expressed by a subjective probability or any other measure of epistemic uncertainty, for example imprecise probabilities. The validation is completed and the model is accredited “good for use”, in the sense that the model output uncertainties are acceptable.

On the contrary, say that the Poisson assumption is hard to justify. Then, the analysts are faced with the choice of arguing for another model or still using the Poisson model as a crude approximation. Suppose the analysts decide for the latter choice, e.g. because they fail to find a rationale for a specific alternative model at this point of the analysis. Then, the model output uncertainty issue is one related to the structural model uncertainty.

To accredit the model, judgments have to made on the magnitude of the error introduced by using the Poisson model. The parameter λ is understood as the average number of events per unit of time. If we know the value of λ, what will be the difference in the true probability distribution of the number of events and in the Poisson distribution, and specifically the difference between the 95% quantilen_{95} corresponding to the true distribution and that of the Poisson distribution? In theory, the analysts can assign a subjective probability distribution reflecting the structural model uncertainty (for different values of λ), but the knowledge that this distribution is based on could be poor and hence the distribution would not provide much insights for concluding on the model being valid or not. Nonetheless, for the sake of example a crude analysis may be conducted to guide the decision on whether or not to accept the model.

To illustrate the analysis consider a case where $n_M = 10$. As explained in Section 2.1, the model error can be written as $\Delta G(\lambda) = G(\lambda) - n_{95} = n_{95}(\lambda) - n_{95}$, where $G(\lambda)$ is the 95^{th}quantileof the Poisson distribution which we refer to as $n_{95}(\lambda)$. The structural model uncertainty relates to uncertainty about the value of $\Delta G(\lambda_0) = G(\lambda_0) - n_{95} = n_{95}(\lambda_0) - n_{95}$, where $\lambda_0$represents the average number of events occurring in [0,t] when considering the infinite population of situations similar to the one studied. Table 1 provides the assigned probabilities for $\Delta G(\lambda_0)$:

Table 1: Assigned Probabilities for $\Delta G(\lambda_0)$:

	(-7.5, -2.5]	(-2.5, 2.5]	(2.5, 7.5]
ity Interval	25%	50%	25%

Whether the results in Table 1 would lead to a recommendation of accreditation or not of the model depends on the situation, *i.e.*, the objective of the analysis for which the model is used and the decision making associated to its outcomes. We would say that in most cases a probability of 25% of an error in the range 2.5-7.5 would be too large relative to the guidance value $n_M = 10$ to justify acceptance.

An alternative approach is to perform qualitative judgments, with the aim of assessing the importance of the structural model error, reflecting how sensitive the produced 95% quantile is with respect to deviations from the underlying distribution and how large the analysts consider the structural model uncertainties. A qualitative score system along these lines can be used, as shown in Flage and Aven [12].

Consider now Example 2 and the model for the number of fatalities N,

$G(X) = 5 X_0 (1- X_1) X_2 X_3 + 10 X_0 (1- X_1) X_2 (1-X_3)$,

This is a physical model of a not yet built plant, and data are not available to check the precision of the model. The suitability of the model has to be determined based on arguments and relevant sub-models. As mentioned above the model is accredited if the analysts find the model producing sufficiently accurate predictions, and to make this assessment the analyst may produce a judgment about the uncertainties on the error G(X)-N.

Different types of judgments may be used, as mentioned for the Poisson example. For example, the analysts may express a subjective probability distribution, based on the available knowledge, of the values of Δ_G or simply an uncertainty interval (e.g. 90%) for Δ_G expressing the degree of confidence of the analysts (e.g. 90%) on the error being within this interval. Depending on the result of this analysis, and the overall purpose of the risk assessment, it may be decided to adjust the model to better reflect the phenomena studied. This adjustment could for example lead to considering the number of fatalities in case of scenarios 2 and 3, to be unknown quantities instead of fixed numbers (5 and 10, respectively). The basis for this model output analysis would of course be an analysis of the structural model uncertainties and the parameter uncertainties. The structural model uncertainty analysis reflects on the accuracy of the number of fatalities predicted by G(X) for specific choices of X (assuming we know the true value of X).

Again qualitative analyses of the importance of the assumptions made should be added to provide a broader basis for making a judgment of the validity of the model.

Similarly, in the Example 3 above concerning the modelling of the groundwater flow in the far field region of a radioactive waste repository site, the four different hypotheses mentioned were considered and accreditation of the models sought using expert evaluation. For each of these four models uncertainty judgments were made on the error Δ_G, using intervals as described above for Example 2. In addition an integrated model were considered [39], based on weights to the four models according to a confidence assignment (the confidence the experts have in the various models for making accurate predictions). The actual weights were 1%, 4%, 35% and 60%. Also uncertainty judgments were made on the error Δ_G for this integrated model.

From the above discussion, it is evident that the issue of accreditation is linked to the question of how to assess the uncertainties about the model difference Δ_G. If multiple experts are engaged in the evaluation and there is no consensus on the hypotheses for modeling the specific phenomena or events of interest, and/or there exist different plausible models, the result could be a wide model difference distribution or interval, which cannot justify model accreditation. For instance, consider the Example 3 above concerning the modelling of the groundwater flow in the far field region of a radioactive waste repository site. Four different hypotheses have been considered and accreditation of the models has been sought by way of their evaluation by experts.

In Zio and Apostolakis [39] and Baraldi and Zio [5], it is shown that the way of handling and aggregating the uncertainties about the difference Δ_G can lead to more or less “spread out” results. Only, if the resulting distribution is judged “sufficiently peaked” or the interval “sufficiently small”, the model can beaccredited. Here again it is clear that arbitrariness in the judgment of the accuracy of the model is a most critical point for the accreditation. The distributions and intervals resulting from the model uncertainty analysis could be judged sufficiently narrow to justify accreditation, in spite that the available evidence is rather vague. Such judgments resulting in over-confidence in the model accuracy can have serious consequences at the decision-making level. For example, the assessment of the hydraulic field in a site proposed to host a radioactive waste repository could be judged adequate based on the results of a porous medium model analysis given few core-boring tests, whereas a fractured medium model would have requested further investigations.

In the strive for accreditation it can be crucial to understand the influence of uncertainties on the results of the analysis, so as to be able to rank their importance for guiding additional efforts (measurements, research, etc.) of uncertainty reduction. This objective is clearly driven by the need of accrediting the model, and enables to build confidence in the model results; but it is a valuable objective to pursue also in cases where model accreditation is not sought. For example, assuming that multiple models are plausible to describe the reality of a situation, given the information available, one may use all models to get insights in the phenomena that govern the situation, within a kind of what–if analysis. Accepting the hypothesis of one particular model, we can study what are the most important uncertainty contributors in this case, and make comparisons among the models. The result is a better understanding of the phenomena and an improved basis for making a decision on the situation at hand, with the required confidence.

If after further understanding and assessment of the assumptions, and possibly a consensus-seeking deliberative process among the multiple experts, the model can still not be accredited, remodelling is required.

In the end, with an accredited model it is possible to perform risk analyses to inform the decision making processes of selection and compliance. To select means to compare performances of alternative system designs, operation modes and maintenance policies, for “optimal" choices. It obviously applies only on accredited models, as confidence is needed for selection. The results of the uncertainty assessments may be expressed in relative, comparative terms on the alternatives under consideration. The confidence needed for this decision task clearly depends on the object of the selection. On the other hand, while the compliance of a system, process and procedure to regulatory criteria applies only to accredited models, the level of confidence needed is absolute. As for selection, it depends on the object of the decision.

For the select and comply objectives, quantification of model output uncertainty is not directly appropriate but there will always be a need for qualitative assessments of model uncertainties as the accredited models will always have limitations and weaknesses, as illustrated by the three examples, and these need to be an integrated part of the results of the risk assessment. A possible way of how to do this qualitative analysis was outlined above for Example 1.

3. Conclusions

The developments of modern science and engineering are strongly reliant on the use of complex models, with increasing demand of sophistication and complexity at all scales. Model predictions are often used to support high-consequence decisions, since simulations are often much less expensive to run than full-scale tests, and in many cases, full-scale tests are not possible at all which is the situation addressed in the present paper.

Given the importance placed on modeling and simulation, a satisfactory level of assurance must be provided that the results obtained from such models are trustworthy for the decision-making purposes for which they are employed. For this, the amount of error and uncertainty that is associated with model predictions must be assessed.

The aim of the present paper has been to contribute to the discussion on how this assessment should be framed and conducted. We believe that new insights have been gained by distinguishing between model output uncertainty, structural model uncertainty and parameter (model input quantities) uncertainty, and making links between the different purposes of modelling and risk assessment. Risk assessments typically address parameter uncertainties but do not cover judgments about the structural model uncertainties. Using some simple examples we have argued that in risk assessment, the issues of analysis of model output uncertainty arise for purposes of accreditation of the models used (verifying that the uncertainties are within limits acceptable for the models’ intended use), whose outcomes inform the decision making about option selection (of system designs, operation modes, *etc.*) and demonstration of criteria compliance. Through the examples we have pointed to the type of logic and argumentation that can lead to accreditation.

For the select and comply objectives, the two last paragraphs of the previous Section provide a summary of our conclusions.

Acknowledgments: The authors are grateful to an anonymous reviewer for his useful comments and suggestions to the original version of this paper.

References

[1]. Apostolakis, G., (1990).*The Concept of Probability in Safety Assessments of Technological Systems.* Science, 250(4986):1359-1364.
[2]. Aven. T. (2011).*Quantitative Risk Assessment:The Scientific Platform.* Cambridge University Press, Cambridge.
[3]. Aven, T. and E. Zio (2011).*Some Considerations on the Treatment of Uncertainties in Risk Assessment for Practical Decision-making.* Reliability Engineering & System Safety 96(1): 64-74.
[4]. Balci, O. and R. Sargent (1982).*Validation of Multivariate Response Simulation Models by using Hotelling'sTwo-sample T2 test.* Simulation, 39(3):185–192.
[5]. Baraldi, P. and E. Zio (2010).*A Comparison between Probabilistic and Dempster-Shafer Theory Approaches to Model Uncertainty Analysis in the Performance Assessment of Radioactive Waste Repositories.* Risk Analysis, 30(7): 1139-1156.
[6]. Bayarri, M. J., J.O.Berger, R. Paulo, J. Sacks, J. A. Cafeo, J. Cavendish, C. H. Lin, and J. Tu.(2007).*A Framework for Validation of Computer Models*.Technometics, 49(2): 138-154.
[7]. Chen, W., L. Baghdasaryan, T.Buranathiti, and J. Cao. (2004*). Model Validation via Uncertainty Propagation*. AIAA Journal, 42(7):1406–1415.
[8]. deRocquigny, E., N. Devictor, and S. Tarantola (eds.) (2008).*Uncertainty in Industrial Practice: A Guide to Quantitative Uncertainty Management*. John Wiley& Sons, N.J.
[9]. Dowding, K., R. G. Hills, I. Leslie, M.Pilch, B.M. Rutherford, and M. L. Hobbs.(2004).*Case Study for Model Validation: Assessing a Model for Thermal Decomposition of Polyurethane Foam.*Technical Report SAND2004-3632, Sandia National Laboratories, Albuquerque, NM.
[10]. Droguett, E.L., and A. Mosleh (2008).*Bayesian Methodology for Model Uncertainty Using Model Performance Data*.Risk Analysis, 28(5):1457-1476.
[11]. Durga Rao K., Kushwaha H. S., VermaA. K. and SrividyaA. (2008)*Epistemic Uncertainty Propagation in Reliability Assessment of Complex Systems*. International Journal of Performability Engineering, 4(1), 71-84.

[12]. Flage, R., and T. Aven. (2009).*Expressing and Communicating Uncertainty in Relation to Quantitative Risk Analysis (QRA).*Reliability: Theory & Applications 2(13): 9-18.
[13]. Hills, R. G.andI. Leslie .(2003). *Statistical Validation of Engineering and Scientific Models: Validation Experiments to Application.* Technical Report SAND2003-0706, Sandia National Laboratories, Albuquerque, NM.
[14]. Jiang, X. and S. Mahadevan.(2007).*Bayesian Risk-based Decision Method for Model Validation under Uncertainty.* Reliability Engineering & System Safety, 92(6):707–718, 2007.
[15]. Jiang, X., R .J. Yang, S.Barbat, andP. Weerappuli. (2009)*Bayesian Probabilistic PCA Approach for Model Validation of Dynamic Systems.*SAE International Journal of Materials and Manufacturing, 2(1): 555-563.
[16]. Kaminski, Jr., J.D.Riera, R.C.R. de Menezes, and L.F.F. Miguel. (2008).*Model Uncertainty in the Assessment of Transmission Line Towers subjected to Cable Rupture*. Engineering Structures 30(10): 2935–2944.
[17]. Kennedy, M. C., and A. O'Hagan.(2001), *Bayesian Calibration of Computer Models*.Journal of the Royal Statistical Society, Series B (Statistical Methodology), 63(3): 425-464.
[18]. Knupp, P., (2002).*Verification of Computer Codes in Computational Science and Engineering*. Chapman & Hall/CRC, Boca Raton, FL.
[19]. Mahadevan, S., and R. Rebba.(2005), *Validation of Reliability Computational Models using Bayes Networks*. Reliability Engineering & System Safety, 87(2):223–232.
[20]. McFarland, J. M., (2008).*Uncertainty Analysis for Computer Simulations through Validation and Calibration*, Ph.D. Thesis, Vanderbilt University, Nashville, U.S.A

[21]. Meeker, W.O., and L.A. Escobar. (1998).*Statistical Methods for Reliability Data*. John Wiley& Sons, N.Y.
[22]. Mosleh, A., N. Siu, C. Smidts, C. Lui (eds.) (1995). *Proceedings of Workshop in Advanced Topicsin Risk and Reliability Analysis. Model Uncertainty: Its Characterization and Quantification.*Center of Reliability Engineering, Univerity of Maryland, U.S.A.
[23]. Nilsen, T., and T. Aven. (2003).*Models and Model Uncertainty in the Context of Risk Analysis*. Reliability Engineering & Systems Safety 79(3): 309-317.
[24]. NUREG- 1855**.** *Guidance on the Treatment of Uncertainties Associated with PRAs in Risk-Informed Decision Making* (NUREG-1855). US-NRC. 2009.
[25]. Oberkampf, W. and Trucano, T., 2002, *Verification and Validation in Computational Fluid Dynamics*. Technical Report SAND2002-0529, Sandia National Laboratories, Albuquerque, New Mexico.
[26]. Oberkampf, W., and T. Trucano. (2006).*Measures of Agreement between Computation and Experiment: Validation Metrics.*Journal of Computational Physics, 217:5–36.
[27]. Oberkampf, W., S. M. DeLand, B.M. Rutherford, K.V. Diegert, and K. F.Alvin.(2002). *Error and Uncertainty in Modeling and Simulation.*Reliability Engineering & Systems Safety 75(3): 333-357.
[28]. Östergaard, C., M. Dogliani, C. Guedes Soares, G.Parmentierand P.T.Pedersen. (1996).*Measures of Model Uncertainty in the Assessment of Primary Stresses in Ship Structures.* Marine Structures. 9(3-4): 427-447
[29]. Paez, T., and A. Urbina. (2002). Validation of Mathematical Models of Complex Structural Dynamic System, In Proceedings of the Ninth International Congress on Sound and Vibration, Orlando, FL
[30]. Parry, G..andM. T. Drouin. (2009).*Risk-Informed Regulatory Decision-Making at the U.S. NRC: Dealing with Model Uncertainty*. Nuclear Regulatory Commission, 2009.
[31]. Rebba, R. *Model Validation and Design under Uncertainty*.Ph.D.Thesis, Vanderbilt University, 2005, Nashville, U.S.A.
[32]. Rebba, R., andS. Mahadevan.(2008).*Computational Methods for Model Reliability Assessment*. Reliability Engineering & System Safety, 93(8):1197-1207.
[33]. Rebba, R., and S. Mahadevan. (2006).*Validation of Models with Multivariate Output.* Reliability Engineering, & System Safety, 91(8):861–871.
[34]. Rebba, R., S. Mahadevan, and S. Huang.(2006).*Validation and Error Estimation of Computational Models.* Reliability Engineering & System Safety, 91(10-11):1390–1397.
[35]. Reinert, J.M., and G. E. Apostolakis. (2006). *Including Model Uncertainty in Risk-informed Decision Making.* Annals of Nuclear Energy, 33(4): 354-369.
[36]. Roache, P., (1998).*Verification and Validation in Computational Science and Engineering*. Hermosa Publishers, Albuquerque, NM, U.S.A.
[37]. Xiong, Y., W. Chen, K. L. Tsui, andD. W.Apley.(2009).*A Better Understanding of Model Updating Strategies in Validating Engineering Models*. Journal of Computer Methods in Applied Mechanics and Engineering, 198(15-16):1327-1337.
[38]. Zio, E., (2006).*A Study of the Bootstrap Method for Estimating the Accuracy of Artificial Neural Networks in Predicting Nuclear Transient Processes*. IEEE Trans. on Nucl. Sci., 53(3): 1460-1478.
[39]. Zio, E., and G. E. Apostolakis (1996).*Two Methods for the Structured Assessment of Model Uncertainty by Experts in Performance Assessments of Radioactive Waste Repositories.*Reliability Engineering & System Safety54(2-3): 225–241.

Terje Aven is Professor of Risk Analysis and Risk Management at the University of Stavanger, Norway. He is also a Principal Researcher at the International Research Institute of Stavanger (IRIS). He is the Vice-Chairman of the European Safety and Reliability Association ESRA. Aven has many years of experience from the petroleum industry (The Norwegian State Oil Company, Statoil). He has published a large number of papers in international journals on probabilistic modelling, reliability, risk and safety. He is the author of several reliability and safety related books, including Stochastic Models in Reliability, Springer Verlag, 1999 (co-author U. Jensen); Foundations of Risk Analysis, Wiley, 2003, 2012; Risk Management, Springer Verlag, 2007 (co-author J.E. Vinnem); Risk Analysis, Wiley 2008; Misconceptions of Risk, Wiley 2010; Risk Management and Governance (co-author O. Renn), Springer Verlag, 2010; and Quantitative Risk Assessment: the Scientific Platform, Cambridge University Press, 2011. He received his Master's degree

(cand.real) and PhD (dr. philos) in Mathematical Statistics (Reliability) at the University of Oslo in 1980 and 1984, respectively.

Enrico Zio is Director of the Chair in Complex Systems and the Energetic Challenge of the European Foundation for New Energy of Electricite’ de France (EDF) atEcoleCentrale Paris and Supelec, full professor, President and Rector's delegate of the Alumni Association and past-Director of the Graduate School at Politecnico di Milano, adjunct professor at University of Stavanger and City University of Hong Kong. He is the Chairman of the European Safety and Reliability Association ESRA and honorary member of the Polish Reliability and Safety Association PSRA, member of the scientific committee of the Accidental Risks Department of the French National Institute for Industrial Environment and Risks, of the Korean Nuclear society and China Prognostics and Health Management society. He is Associate Editor of IEEE Transactions on Reliability and editorial board member in various international scientific journals. He has functioned as Scientific Chairman of three International Conferences and as Associate General Chairman of three others. His research focuses on the characterization and modeling of the failure/repair/maintenance behavior of components, complex systems and critical infrastructures for the study of their reliability, availability, maintainability, prognostics, safety, vulnerability and security, mostly using a computational approach based on advanced Monte Carlo simulation methods, soft computing techniques and optimization heuristics. He is author or co-author of five international books and more than 200 papers on international journals.