ESTIMATION OF UNCERTAINTY IN RISK ASSESSMENT OF HYDROGEN APPLICATIONS

Markert, F.[1], Krymsky, V.[2], and Kozine, I.[1]
[1] Safety, Reliability & Human Factors section, Technical University of Denmark, Management Engineering Department, Produktionstorvet building 426, DK-2800 Kongens Lyngby, Denmark fram@man.dtu.dk
[2] Ufa State Academy of Economics and Service, 145 Chernyshevsky Street, Ufa, Bashkortostan, 450078, Russia vikrymsky@mail.ru

ABSTRACT

Hydrogen technologies such as hydrogen fuelled vehicles and refuelling stations are being tested in practice in a number of projects (e.g. HyFleet-Cute and Whistler project) giving valuable information on the reliability and maintenance requirements. In order to establish refuelling stations the permitting authorities request qualitative and quantitative risk assessments to show the safety and acceptability in terms of failure frequencies and respective consequences. For new technologies not all statistical data can be established or are available in good quality causing assumptions and extrapolations to be made. Therefore, the risk assessment results contain varying degrees of uncertainty as some components are well established while others are not. The paper describes a methodology to evaluate the degree of uncertainty in data for hydrogen applications based on the bias concept of the total probability and the NUSAP concept to quantify uncertainties of new not fully qualified hydrogen technologies and implications to risk management.

1. NOMENCLATURE

θ_i	State of the technological process
$\Pr(\theta_i)$	Probability of being in a state θ_i of the technological process
X	Consequence of the hazardous state
x_j	Discretised value of the consequence measure
x^*	Critical value of X within a particular time period
$F(x^*)$	Probability that the consequence does not exceed the critical value x^*
R	Risk measure
Δ	Bias (difference between the true and estimated value of a risk measure)
$Sc_i^{(j)}$	Score provided by the j-th expert when answering the i-th question in the checklist
$\gamma^{(j)}$	Analogue of a 'degree of belief' of expert computed via $Sc_i^{(j)}$
γ	Combined 'degree of belief' of the expert group
w_j	"Weight" of an expert reflecting the quality of individual judgement
β_j	Measure of difference between the scores given by the definite expert and 'true' scores for the proven technology ('reference point')

2. INTRODUCTION

Development, investigation and release of the emerging hydrogen technology, as any other technology, requires many decisions and actions that aim at accomplishing societal acceptance criteria for the specific technology. In this context, it is important to be precautious with regard to prevention of accidents or catastrophes caused by the side effects or undesirable factors not taken into account. Each technology typically starts from a basic idea, which may be very similar to former ones, but of course not fully in order to provide the progress in achieving one or more goals. So implementation of an emerging technology into a new socio-technological system, always will introduce some degree of uncertainty due to lack of knowledge. Therefore, the implementation of the new technology is normally being done via a step-by-step procedure, which in particular supposes theoretical research as well as different tests/ verification.

The amount of knowledge about a technology performance naturally increases while the process moves from one step to another. The initial steps are accompanied by a very high level of uncertainty because of the lack of knowledge, absence of experience and statistical data. Further steps throw light on the originally unknown (or imperfectly known) aspects of technology influence on people, environment, assets, etc. The level of uncertainty reduces due to the supplementary information brought from various sources, e.g. experiments, expert judgments or theoretical insights. The right moment for release of the new technology can be defined as the time at which our knowledge and precision of our forecast are good enough to guarantee the “acceptable risk”, which is a certain risk level not in contradiction with societal expectations. A framework for such a technology qualification is described by Det Norske Veritas [1,2].

During the last decade, a number of hydrogen technologies have been designed and developed to a level suitable for marked introduction. The car industries have created hydrogen fueled cars based on combustion or fuel cell technologies and in parallel developed the required infrastructures [3] such as refueling stations. In order to achieve this, new components as e.g. storage vessels utilizing 700 bar pressure technology and isolating tanks to store liquid hydrogen have also been developed. The feasibility of these new technologies is being tested in large scale projects as e.g. the European HyFleet-Cute [4,5] and “Hydrogen highway” projects[*], as e.g. the American Whistler project to give green 2010 Olympic winter games. In order to address the lack of knowledge for the emerging hydrogen technologies and their societal impact the NoE HySafe [6] was funded by the EU and the International Energy Agency Hydrogen Implementing Agreement established complementary activities within their “task 19 “Hydrogen Safety activities” [7].

“The prioritization of the HySafe internal project activities was based on a phenomena identification and ranking exercise (PIRT) and expert interviews. The identified research headlines were “Releases in (partially) confined areas”, “Mitigation” and “Quantitative Risk Assessment”. Along these headlines existing or planned research work was re-orientated and slightly modified, to build up three large internal research projects “InsHyde”, “HyTunnel”, and “HyQRA [8]*”. In InsHyde realistic indoor hydrogen leaks and associated hazards have been investigated to provide recommendations for the safe use of indoor hydrogen systems including mitigation and detection means. The appropriateness of available regulations, codes and standards (RCS) has been assessed. Experimental and numerical work was conducted to benchmark simulation tools and to evaluate the related recommendations* [9-11]*. HyTunnel contributed to the understanding of the nature of the hazards posed by hydrogen vehicles inside tunnels and its relative severity compared to other fuels. In HyQRA quantitative risk assessment strategies were applied to relevant scenarios in a hydrogen refueling station and the performance was compared to derive also recommendations”.*

The IEA HIA Task 19 efforts are e.g. to develop guidelines and criteria for evaluation purposes within Quantitative Risk Assessment (QRA) studies of hydrogen facilities [12-14]. These QRAs are needed for the permitting process e.g. for hydrogen refueling stations to validate the required safety performance. There have been a number of other studies stressing the importance of access to the relevant accident and incident information, which is still limited for hydrogen-specific incidents and accidents [15,16]. The authors discuss the uncertainty contained in RA that is inherent to the methodology due to both aleatory and epistemic uncertainty[†] and also refer to former benchmark exercises [17-19] trying to quantify these uncertainties. The lack of reliability data for hydrogen technologies may be e.g. solved by using data for comparable hydrocarbon incidents [20], instead. This, of course, generates a potential source of inaccuracy in QRA studies, as the quality of the results depend strongly on the data reliability and quality.

[*] http://en.wikipedia.org/wiki/Hydrogen_highway

[†] Aleatory uncertainty is often called stochastic or irreducible while epistemic uncertainty is called reducible which stems from a lack of knowledge.

Partially, audit techniques may be helpful to improve the quality of management systems to be established which are concerned with e.g. the reliability and effectiveness of safety barriers and the probability of recognized incident and accident scenarios [21]. In order to avoid ambiguity in the expert communications, it is further important to regard the terminology [22] as a possible source of uncertainty:*"The sciences analyzing and describing risks are relatively new and developing, and the associated terminologies are developing as well. This has led to ambiguity in the use of terms, both between different risk sciences and between the different parties involved in risk debates."*

As stated above, uncertainty is normally differentiated into aleatory uncertainty and epistemic uncertainty. The former is represented by statistical measures, as e.g. variations in wind speeds. The latter covers uncertainties due to incomplete knowledge as e.g. known unknowns and unknown unknowns. Paté-Cornell [23] identified six levels of uncertainty treatment:

- Level 0: Simply procedure of hazard detection and failure modes identification
- Level 1: "Worst case" approach
- Level 2: "Quasi-worst cases" and plausible upper bounds
- Level 3: Best estimates and central values
- Level 4: Probabilistic risk assessment, single risk curve
- Level 5: Probabilistic risk analysis, multiple risk curves

Paté-Cornell argues that all levels have their role and it is not always demanded to use level 4 or 5 assessments to describe the uncertainties, as in simple cases with low cost solutions a level 0 approach may be fully appropriate. In the following part of the article, a methodology for uncertainty assessment is described that can be used to evaluate and quantify the uncertainty in RA studies of emerging hydrogen technologies using expert judgements.

3. RISK MEASURE AND GENERAL UNCERTAINTY MODEL

The work described in the current paper aims at attempting to derive a procedure which can be used both at earlier and later phases of the technology development as described in a report by Krymsky [24]. The procedure suggests the quantification of uncertainties even if they are preliminarily characterized qualitatively. The next sections outline briefly how this quantification can be done and an example shows the applicability of it within RA of refueling stations.

The model we apply to solving our problem operates with the totality of discrete and mutually exclusive states θ_i of the technological process. Each state is associated with a probability $\Pr(\theta_i)$. Furthermore, each hazardous state is associated with consequences X, which can also be discretised into consequences x_j, each associated with a probability for each state of the technological process $\Pr(x_j|\theta_i)$. For instance, to estimate the risk of harm to people resulting from a flammable gaseous cloud within a certain area (consequence) as the result of a gaseous release (process state), we need i.a. to discretise the amount of the released substances and specific weather conditions represented by certain wind speeds, wind directions, rain/ no rain, substance type, its properties, etc., and to assign a probability for their combined occurrences. After that we need to assign probabilities to the occurrence of certain defined end points (e.g. harm to people) to evaluate the consequences, given the particular release and weather conditions.

We choose the risk measure $R(x^*)$ as a probability of the consequences with the level greater or equal to the critical (societally acceptable) value x^* within a particular operating timeframe (e.g. one calendar year). It completely satisfies the requirements of the International Standard ISO 31000, namely risk measure is expressed in terms of combination of consequences of an event and associated likelihood of occurrence;

The specific point of our approach is that expression for the formulated risk measure $R(x^*)$ takes the form [25-27]:

$$R = R(x^*) = \Pr(X \geq x^*) = \sum_{\forall j:\, x_j \geq x^*} \sum_{i=1}^{n} P(x_j / \theta_i) \cdot P(\theta_i) + \Delta,$$

where, Δ is termed the '*bias*' and $\sum_{i=1}^{n} P(\theta_i)$.

Actually, the bias is an additive term in the expression for computing the risk measure to take into account the uncertainty which accompanies the accomplishment of our risk assessment algorithm [25-27]. In practice the value $P_1 = \sum_{\forall j:\, x_j \geq x^*} \sum_{i=1}^{n} P(x_j / \theta_i) \cdot P(\theta_i)$ obtained from direct computations differs from the true value of the risk measure R namely due to the uncertainty contributions. The bias $\Delta = R - P_1$.is the difference of the two probabilities, so it is dimensionless and Δ is considered as a random value [25-27].

Actually, the bias is produced by our doubts on the correctness/ precision of the model we use for risk assessment. We cannot be sure that: a) the list of the process states is complete; b) the probabilities of the states are precise; c) the conditional probabilities 'consequence conditioned on the definite state' are precise. Such probabilities are frequently elicited from the experts dealing with emerging technology. So in fact, we have to deal with *subjective* probabilities which depend on the expert personalities and questionnaires/ procedures. The other group of factors, which influences the quality of the model, is concerned with discretising the totality of the states and the set of consequence values. Such an approach necessarily leads to approximate description of the risk dependency on accident likelihood and correspondent consequences. However, the degree of the discretisation effect cannot be presented in deterministic codes. Summing up we can say that:

(i) the risk is computed via the model based on the formula of the total probability $P_1 = \sum_{\forall j:\, x_j \geq x^*} \sum_{i=1}^{n} P(x_j / \theta_i) \cdot P(\theta_i),\ \sum_{i=1}^{n} P(\theta_i)$

this model captures aleatory uncertainty associated with the scenarios of accidents;

(ii) any model used for risk assessment is not perfect, this fact causes the appearance of the bias term which captures epistemic uncertainty.

The specific feature of a new technology qualification problem is the high importance of the epistemic uncertainty, because this reduces the belief in the adequacy of the computed risk using traditional computations. The bias is formed by contributions of different types of uncertainty within the classical point of view [25] and in the framework of imprecise hierarchical models [28].

The next section outlines the mechanism proposed for creating the model which describes the interconnection between epistemic uncertainty and the bias as a term in the expression for risk measure.

MODEL FOR QUANTIFICATION OF EPISTEMIC UNCERTAINTY

The new technology has to be accompanied by risk assessments from the initial stages of its development. Naturally the very early stages of technology development are performed with a lack of reliable and credible information. So for technology qualification we need special tools which are convenient when operating with unreliable data.

A detailed survey of the tools for uncertainty assessment is presented in the report [29]. It covers:

• Sensitivity Analysis (screening, local, global);
• Error propagation equations (“Tier 1”);
• Monte Carlo Analysis (“Tier 2”);
• Expert Elicitation;
• NUSAP (Numeral -Unit -Spread -Assessment -Pedigree);
• Scenario Analysis;
• PRIMA (Pluralistic fRamework of Integrated uncertainty Management and risk Analysis);
• Checklist for Model Quality Assistance.

Analyzing the different approaches, we apply the NUSAP method [30] to quantify the uncertainty. This method combining quantitative and qualitative uncertainty measures is developed for policy purposes [29] and it ‘fosters an enhanced appreciation of the quality of information’. This is very important when the amount of the reliable information is limited.

Van der Sluijs et al. [31] give the explanation of the NUSAP components: The last two qualifiers constitute the more qualitative side of the NUSAP expression, which is of importance for us.

‘Numeral; this will be an ordinary number, but when appropriate it can be a more general quantity, such as the expression ‘a million’.
Unit, this may be just of ‘units’ to qualify the numeral, but may also contain extra information on the significant digits in the numeral, as e.g expressed by a dimension at which the unit is evaluated (e.g. 10^2kg).
Spread generalizes from ‘random error’ of experiments or ‘variance’ of statistics. Methods to address Spread can be statistical data analysis, sensitivity analysis, or Monte Carlo analysis, possibly in combination with expert elicitation.
Assessment expresses judgments about the information. In the case of statistical tests, this might be of a significance level; in the case of numerical estimates for policy purposes, it might be the qualifier ‘optimistic’ or ‘pessimistic’.
Pedigree conveys an evaluative account of the production process of information, and indicates different aspects of the underpinning of the numbers and scientific status of the knowledge used.

Pedigree is expressed by means of a set of pedigree criteria to assess these different aspects. Assessment of pedigree involves qualitative expert judgment. To minimize arbitrariness and subjectivity in measuring strength, a pedigree matrix is used to code qualitative expert judgments for each criterion into a discrete numeral scale from 0 (weak) to 4 (strong) with linguistic descriptions (modes) of each level on the scale. Each special sort of information has its own aspects that are key to its pedigree; so different pedigree matrices using different pedigree criteria can be used to qualify different sorts of information’. The example of pedigree matrix brought from paper [32] and addressing the quality of data which characterize radiological model parameters is presented in Table 1.

Table 1 Example of a pedigree matrix [32]

Type	**Source**	**Relevance**	**Processing**	**Code**
Constants	Reviewed	Full	Confirmed	4
Deduced	Refereed	High	Aggregated	3
Estimated	Internal	Good	Extended	2
Synthesised	Conference	Medium	Accepted	1
Hypothetical	Isolated	Poor	Copied	0

Of course, such a strategy to analyse the information requires special checklists which contain relevant questionnaires for experts. For instance, the paper [33] gives one of the possible checklist versions applied to assessing waterborne risks indicated in Table 3 (see in appendix). It has been based on the scores derived from brainstorming sessions among a small group of experts.

Now we propose the procedure aiming at forming decision support on the basis of obtained scores. We start from the consideration of individual expert judgment.

Assume that the checklist contains *N* rows with the questions. Each *i*-th question will be answered by *j*-th expert with the score $Sc_i^{(j)} \in [0,4]$.

Obviously, $\min \sum_{i=1}^{N} Sc_i^{(j)} = 0;\ \max \sum_{i=1}^{N} Sc_i^{(j)} = 4N.$

We can compute the j-th expert's 'degree of belief' $\gamma^{(j)}$ in the precision of the value P_1 of the basic model of a specific risk assessment, which satisfies

$$0 \le \gamma^{(j)} = \left(\sum_{i=1}^{N} Sc_i^{(j)} \right) / (4N) \le 1.$$

Actually, $\gamma^{(j)}$ is:

$$P_1 = \sum_{\forall j: x_j \ge x^*} \sum_{i=1}^{n} P(x_j / \theta_i) \cdot P(\theta_i) .$$

So, it can be considered as some analogue to a subjective probability. The next step should be the aggregation of the individual judgments, as we compute the value of

$$\gamma = \sum_{j=1}^{K} w_j \gamma^{(j)},$$

Where, γ is the combined 'degree of belief' of the expert group in the quality of risk assessments; *K* is the number of experts in the group, w_j is the weight associated with *j*-th expert.

Naturally, w_j must reflect the quality of *j*-th individual judgement, which requires a 'reference point', which could be any proven technology treated via NUSAP. These results are assumed to be the known 'true' scores, $Sc_i^{ref}, i = 1,2,...,N$. Simultaneously, the *j*-th expert being asked the same questions in relation to the reference point would give the scores $Sc_i^{(j)ref}, i = 1,2,...,N$. It is easy to see that: $\max \sum_{i=1}^{N} \left(Sc_i^{(j)ref} - Sc_i^{ref} \right)^2 = 16N.$

Hence, $w_j = \beta_j / \sum_{r=1}^{K} \beta_r$, where $\beta_j = 16N - \sum_{i=1}^{N} \left(Sc_i^{(j)ref} - Sc_i^{ref} \right)^2$,

just satisfies the typical requirements to the weight coefficients:

$$w_j \ge 0,\ \sum_{j=1}^{K} w_j = 1.$$

Besides, the growth of w_j corresponds to the fact that the values of $Sc_i^{(j)ref}$ and Sc_i^{ref} are getting tighter.

Now, we have an opportunity to come back to our basic model for risk assessment and consider the dependency of the bias in terms of γ. It is clear when $\gamma \to 1$ then the width of the bias reduces to zero (as we have no doubts in the quality of risk assessments). On the contrary when $\gamma \to 0$ then the width of the bias increases to its maximum value 1 (as there is no belief at all in the quality of assessments).

The total interval for the bias consists of the two subintervals:

A 'negative' $\left[-\sum_{\forall j:\, x_j \ge x^*} \sum_{i=1}^{n} P(x_j/\theta_i)\cdot P(\theta_i);\, 0\right)$ and

a 'positive' $\left[0;\ 1-\sum_{\forall j:\, x_j \ge x^*} \sum_{i=1}^{n} P(x_j/\theta_i)\cdot P(\theta_i)\right]$.

For the 'negative' subinterval we can compute a modified estimation of its width which takes into account the results of NUSAP procedure application:

$$\Delta_{Sc}^{-} = \left(\sum_{\forall j:\, x_j \ge x^*} \sum_{i=1}^{n} P(x_j/\theta_i)\cdot P(\theta_i)\right)(1-\gamma).$$

The same way of thinking for the 'positive' bias subinterval brings us to the expression:

$$\Delta_{Sc}^{+} = \left(1-\sum_{\forall j:\, x_j \ge x^*} \sum_{i=1}^{n} P(x_j/\theta_i)\cdot P(\theta_i)\right)(1-\gamma).$$

As a result we conclude that the growth of γ (the 'degree of belief' in the correctness of risk assessments) decrease the total width for the bias interval. At the initial stages of the technology development (the stages of very restricted knowledge about its performances) γ is small, and consequently the bias is large. After some time we can involve additional knowledge based on theoretical insights and practical investigations (experiments, exploitation); naturally, γ becomes larger. In the limiting case, when $\gamma \to 1$, the bias approaches zero indicating that the experts have no doubts in the quality of the obtained risk assessments. This would be the case when the epistemic uncertainty completely disappears and only the aleatory uncertainty is left in the results of the risk assessment.

AN EXAMPLE

Table 2 Risk model input data and assumption taken from [33]

Item	**Scenario numbers and descriptions**	**Release hole size (mm)**	**Max quality released (kg)**	**Initial failure frequency**	**Usage frequency**	**Total release frequency**
Release pressure 160 bar						
Tube trailer (8 tubes, 24 fittings in total)	1 Catastrophic failure of tube	N/A	26.8	1×10^{-6}/item/yr	Always in pressurized condition	8×10^{-6}/trailer/yr
	2 Leak from tube trailer fittings (10 mm)	10	241	1×10^{-5}/item/yr		2.4×10^{-4}/trailer/yr
	3 Full bore rupture of flexible hose from tube	13	241	4×10^{-6}/h	1 h in total/per day for two hoses	7.3×10^{-4}/hose/yr
Pipe work-1 to	4 Full bore rupture of	21	241	1×10^{-6}/m/yr	Always in	1.7×10^{-5}/yr

compressors (17 m long, 3 m height)	pipe work				pressurized condition	
Release pressure 414 bar						
Two compressors with one standby	5 Catastrophic failure of compressor	N/A	2	6.5×10^{-3}/item/yr	Time fraction 50% for one and 25% for the other	4.875×10^{-3}/yr
	6 Leak from compressor	10	243	5.85×10^{-2}/item/yr		4.3875×10^{-2}/yr
Release pressure 390 bar						
Buffer storage (9 tubes, 34 fittings in total)	7 Catastrophic failure of storage tube	N/A	19.5	1×10^{-6}/item/yr	Always in pressurized condition	9×10^{-6}/yr
	8 Leak from buffer storage fittings	10	175.5	1×10^{-5}/item/yr		3.4×10^{-4}/yr
Pipe work-2 to dispenser (16 m long, in trench)	9 Full bore rupture of pipe work	21	175.5	1×10^{-6}/m/yr	Always in pressurized condition	1.6×10^{-5}/yr
Dispenser (*Release pressure 350 bar*)	10 Catastrophic failure of dispenser (including flexible hose rupture)	N/A	1	4×10^{-6}/h	1 h/per day	1.46×10^{-3}/yr

The comments to Table 2 taken from [34] are as following:

"Risk contribution analysis shows that the compressor leak contributes most to the individual risk of two vulnerable spots inside the station: the center of the control room and the refueling spot near the dispenser. These two sites are the places where workers most frequently go or stay. It can be seen from Table 2 that compressor leak contributes 99% and 68% to the total individual risk of the control room center and the refueling spot, respectively." For the scenario "the individual risk at the center of the control room" Zhiyong Li et al. [33] calculated a total individual risk of 3.42 x 10^{-4}.

Now, let us look at these assessments from the point of view on 'new technology qualification' approach using the following hypothetical arguments and three artificial experts. Let us assume that the data used for the total frequency estimations have some unknown degrees of uncertainty and thus are not completely reliable and credible. The sources of the information chosen by the authors of the paper [34] are well-known documents (TNO Purple Book and HSE Report on Hydrogen Releases Statistics, 2001), which contain the data based on the proven assumptions and real exploitation experience. However, they both are issued about 10 years ago and do not take the statistics of the recent accidents / failures into account. Besides, they principally do not deal with any specific features of hydrogen equipment exploitation for the conditions found in the Shanghai region (in particular, weather conditions, e.g. temperature, humidity, air pressure, etc.). Consequently, the results of the risk assessment are not precise (due to the influence of the second order uncertainty), and each of the failure frequencies given in the Table 2 also should be characterized by a specific bias.

According to the thoughts described before, we may conclude that the bias related to the total individual risk for the personnel in the centre of the control room can be estimated as an interval which consists of a 'negative' $[-3.42\times10^{-4}; 0)$ and a 'positive' $[0;\ 1-3.42\times10^{-4}]$ part.

The next step is applying NUSAP-based methodology for computing corrected bounds for the bias subintervals. Assume we have 3 experts who answered the questionnaire as presented in Table 3 (Appendix) giving their judgments on the quality of the numerical data. Note that the number of questions here is N=12 and the max score is 4N=48. The results of the expert procedure are the following:

$$\sum_{i=1}^{12} Sc_i^{(1')} = 42;\ \sum_{i=1}^{12} Sc_i^{(2)} = 44;\ \sum_{i=1}^{12} Sc_i^{(3')} = 38.$$

Consequently

$\gamma^{(1)} = \left(\sum_{i=1}^{12} Sc_i^{(1)} \right) / 48 = 0.875000;$ $\gamma^{(2)} = \left(\sum_{i=1}^{12} Sc_i^{(2)} \right) / 48 = 0.897959;$

$\gamma^{(3)} = \left(\sum_{i=1}^{12} Sc_i^{(3)} \right) / 48 = 0.791667.$

Let us consider the phase which allows aggregating the individual expert judgments. Assume the comparison of judgment quality in relation to the 'reference point', which lead to the 'weight for each expert': w_1=0.629; w_2=0.274; w_3=0.097; Summation results in the following degrees of believe:

$$\gamma = \sum_{j=1}^{3} w_j \gamma^{(j)} = 0.873207; \qquad 1 - \gamma = 0.126793.$$

If we now multiply the bounds for the bias subintervals by 0.126793, we obtain the following corrected estimations: [-0.043 x 10^{-4}; 0) and [0, 0.12675]

It is easy to see that the interval for the bias becomes much tighter. However the influence of uncertainty remains very essential, so we cannot be sure that the individual risk equals IR = 3.42x10^{-4}. There is a chance that it may be much higher because of the estimated bias, possibly as high as

IR = 3.42 x 10^{-4} + 0.12675 = 0.12709.

This hypothetical example revealed a large contribution of epistemic uncertainty in the results of the risk assessment. There may be various sources of uncertainty present starting with the imperfect knowledge on the technology and the usage of not fully validated reliability data used for geographic regions with different climate zones to the individual knowledge of the experts performing RA. The conclusions could be to continue to gain knowledge about the new technology, to collect additional validated data and repeating the qualification procedure in order to find more justified risk assessments.

4. CONCLUSIONS

The article presented the role of uncertainty that is identified in the emerging hydrogen technologies and provided a brief review on the current development in this field. Further to enable the quantification of the uncertainty, a specific procedure has been described combining different concepts described in the literature. The procedure for the new technology qualification proposed is based on the combination of two principal parts:

1. The notion of the total risk measure including the so called 'bias'[25];

2. NUSAP approach to fulfilling the operations on the numerical data in the conditions of high uncertainty [30].

This procedure for the qualification of new and emerging technologies allows analyzing the hazards associated with the above technology even for a high level of uncertainty. Actually, we become able to apply any known methodology of risk assessment taking into account imprecision of the results due to incomplete reliability and credibility of the initial data and the model as well. Naturally, this brings us to the implementation of the 'precautionary principle': if epistemic uncertainty influences too much, we have to be ready to deal with the worst scenario (when the real risk measure equals the assessed one plus bias).

However step-by-step investigation of the technology gives an opportunity to reduce the interval for the bias, consequently the risk assessments become more precise. This process could be an implementing mechanism for choosing the preferable time for releasing a new technology. Such a time would correspond to the stage were the new technology would meet the acceptable risk level and simultaneously it provides some evidence that the above assessments indeed are adequate (very close to their real values). The article outlines one of the ways of quantifying the uncertainty for the purposes of preventing the undesirable impacts of the new technologies on people and environment. Meanwhile, it can be the basis for future research with an increased number of the case studies. For risk management purposes the acceptance criteria could be extended to include both the severity and the uncertainty, as it is suggested by some authors, e.g. based on the degree of uncertainty in the RA, Renn [35-38] suggested to apply different risk management strategies for these emerging technologies, e.g. QRA based, precautionary or informative approach.

REFERENCES

[1] Qualification procedures for New Technologies, Recommended Practice, DNV-RP-A203 (2001) 1-42.

[2] Technology Qualification Management, Offshore Service Specification, DNV-OSS-401 (2006) 1-24.

[3] F. Markert, S.K. Nielsen, J.L. Paulsen, V. Andersen, Safety aspects of future infrastructure scenarios with hydrogen refuelling stations, Int J Hydrogen Energy. 32 (2007) 2227-2234.

[4] K. Haraldsson, A. Folkesson, M. Saxe, P. Alvfors, A first report on the attitude towards hydrogen fuel cell buses in Stockholm, Int J Hydrogen Energy. 31 (2006) 317-325.

[5] M. Saxe, A. Folkesson, P. Alvfors, A follow-up and conclusive report on the attitude towards hydrogen fuel cell buses in the CUTE project--From passengers in Stockholm to bus operators in Europe, International Journal of Hydrogen Energy. 32 (2007) 4295-4305.

[6] T. Jordan, et al. Achievements of the EC network of excellence HySafe, Int J Hydrogen Energy. 36 (2011) 2656-2665.

[7] A. Tchouvelev, Knowledge Gaps in Hydrogen Safety, R1_Jan-08 1-54.

[8] K. Ham, A. Marangon, P. Middha, N. Versloot, N. Rosmuller, M. Carcassi, O.R. Hansen, M. Schiavetti, E. Papanikolaou, A. Venetsanos, A. Engebø, J.L. Saw, J. Saffers, A. Flores, D. Serbanescu, Benchmark exercise on risk assessment methods applied to a virtual hydrogen refuelling station, Int J Hydrogen Energy. 36 (2011) 2666-2677.

[9] E. Gallego, E. Migoya, J.M. Martin-Valdepenas, A. Crespo, J. Garcia, A. Venetsanos, E. Papanikolaou, S. Kumar, E. Studer, Y. Dagba, T. Jordan, W. Jahn, S. Hoiset, D. Makarov, J. Piechna, An intercomparison exercise on the capabilities of CFD models to predict distribution and mixing of H2 in a closed vessel, International Journal of Hydrogen Energy. 32 (2007) 2235-2245.

[10] A.G. Venetsanos, E. Papanikolaou, M. Delichatsios, J. Garcia, O.R. Hansen, M. Heitsch, A. Huser, W. Jahn, T. Jordan, J.M. Lacome, H.S. Ledin, D. Makarov, P. Middha, E. Studer, A.V. Tchouvelev, A. Teodorczyk, F. Verbecke, d.V. Van, An inter-comparison exercise on the capabilities of CFD models to predict the short and long term distribution and mixing of hydrogen in a garage, International Journal of Hydrogen Energy. 34 (2009) 5912-5923.

[11] A.G. Venetsanos, P. Adams, I. Azkarate, A. Bengaouer, L. Brett, M.N. Carcassi, A. Engebø, E. Gallego, A.I. Gavrikov, O.R. Hansen, S. Hawksworth, T. Jordan, A. Kessler, S. Kumar, V. Molkov, S. Nilsen, E. Reinecke, M. Stöcklin, U. Schmidtchen, A. Teodorczyk, D. Tigreat, N.H.A. Versloot, On the use of hydrogen in confined spaces: Results from the internal project InsHyde, Int J Hydrogen Energy. 36 (2011) 2693-2699.

[12] J. LaChance, A. Tchouvelev, A. Engebo, Development of uniform harm criteria for use in quantitative risk analysis of the hydrogen infrastructure, Int J Hydrogen Energy. 36 (2011) 2381-2388.

[13] J. LaChance, Risk-informed separation distances for hydrogen refueling stations, International Journal of Hydrogen Energy. 34 (2009) 5838-5845.

[14] J. LaChance, A. Tchouvelev, J. Ohi, Risk-informed process and tools for permitting hydrogen fueling stations, International Journal of Hydrogen Energy. 34 (2009) 5855-5861.

[15] H.J. Pasman, W.J. Rogers, Safety challenges in view of the upcoming hydrogen economy: An overview, J Loss Prev Process Ind. 23 (2010) 697-704.

[16] H.J. Pasman, Challenges to improve confidence level of risk assessment of hydrogen technologies, Int J Hydrogen Energy. 36 (2011) 2407-2413.

[17] A. Amendola, Uncertainties in systems reliability modelling: Insight gained through European Benchmark exercises, Nucl. Eng. Des. 93 (1986) 215-225.
[18] A. Amendola, S. Contini, I. Ziomas, Uncertainties in chemical risk assessment: Results of a European benchmark exercise, J. Hazard. Mater. 29 (1992) 347-363.
[19] K. Lauridsen, I.O. Kozine, F. Markert, A. Amendola, M. Christou, M. Fiori, Assessment of Uncertainties in Risk Analysis of Chemical Establishments. The ASSURANCE project. Risø-R-1344(EN) (2002) 1-52.
[20] G.P. Haugom, P. Friis-Hansen, Risk modelling of a hydrogen refuelling station using Bayesian network, Int J Hydrogen Energy. 36 (2011) 2389-2397.
[21] F. Guldenmund, A. Hale, L. Goossens, J. Betten, N.J. Duijm, The development of an audit technique to assess the quality of safety barrier management, Journal of Hazardous Materials. 130 (2006) 234-241.
[22] F.M. Christensen, O. Andersen, N.J. Duijm, P. Harremoës, Risk terminology—a platform for common understanding and better communication, J. Hazard. Mater. 103 (2003) 181-203.
[23] M.E. Paté-Cornell, Uncertainties in risk analysis: Six levels of treatment, Reliab. Eng. Syst. Saf. 54 (1996) 95-111.
[24] V.G. Krymsky, Risk assessment and new technology qualification: a novel framework, Report of DTU Guest Researcher, Management Eng. Department, DTU, Denmark (2011) 1-21.
[25] E. Zio, G.E. Apostolakis, Two methods for the structured assessment of model uncertainty by experts in performance assessments of radioactive waste repositories, Reliab. Eng. Syst. Saf. 54 (1996) 225-241.
[26] T. Nilsen, T. Aven, Models and model uncertainty in the context of risk analysis, Reliab. Eng. Syst. Saf. 79 (2003) 309-317.
[27] T. Rosqvist, R. Tuominen, Qualification of Formal Safety Assessment: an exploratory study, Saf. Sci. 42 (2004) 99-120.
[28] I.O. Kozine, L.V. Utkin, Processing unreliable judgements with an imprecise hierarchical model, Risk Decis. Policy. 7 (2002) 325 -339.
[29] J.P. van der Sluijs, P.H.M. Janssen, A.C. Petersen, P. Kloprogge, J.S. Risbey, W. Tuinstra, J.R. Ravetz, RIVM/MNP Guidance for Uncertainty Assessment and Communication: Tool Catalogue for Uncertainty Assessment, Report nr: NWS-E-2004-37, Copernicus Institute & RIVM; Utrecht/Bilthoven, The Netherlands (2004) 1- 60.
[30] S.O. Funtowicz, J.R. Ravetz, Uncertainty and Quality in Science for Policy, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1990.
[31] J.P. Van Der Sluijs, M. Craye, S. Funtowicz, P. Kloprogge, J. Ravetz, J. Risbey, Combining Quantitative and Qualitative Measures of Uncertainty in Model-Based Environmental Assessment: The NUSAP System, Risk Analysis. 25 (2005) 481-492.
[32] S.O. Funtowicz et al.; Quality assessment for radiological model parameters; Journal of Radiological Protection.; 263.
[33] S.M. Macgill, L. Fewtrell, D. Kay, Towards quality assurance of assessed waterborne risks, Water Res. 34 (2000) 1050-1056.
[34] L. Zhiyong, P. Xiangmin, M. Jianxin, Quantitative risk assessment on a gaseous hydrogen refueling station in Shanghai, Int J Hydrogen Energy. 35 (2010) 6822-6829.
[35] O. Renn, Risk Governance - Coping with Uncertainty in a Complex World, Earthscan, London, Washington, DC, 2008.
[36] O. Renn, A. Klinke, Systemic risks: a new challenge for risk management, EMBO reports. 5 (2004) S41-S46.
[37] O. Renn, Risks and Society, International Conference: Risk Analysis and Its Role in the European Union, Brussels, Belgium (2000).
[38] O. Renn, Emerging risks in public perception: will we face an acceptance crisis? 1st iNTeg-Risk conference, Stuttgart, Germany (2009).

APPENDIX

Table 3 Example for a quality audit checklist and scoring framework taken from [33]

Dimension	Question	Level	Score
Observation			
Measure	How close a match is there between what is being observed and the measure adopted to observe it?	Primary Standard Convenience Symbolic (…)	4 3 2 1
Data	How strong is the empirical content?	Bespoke Direct Calculated Educated guess (…)	4 3 2 1
Sensitivity	How critical is the measure to the stability of the result?	Strong Resilient Variable (…)	4 3 2
Method			
Theory	How strong is the theoretical base?	Laws Well-tested theories Emerging theories/computational models Hypothesis/statistical processing (…)	4 3 2 1
Robustness	How robust is the result to changes in methodological specification?	Strong Resilient Variable (…)	4 3 2
Output			
Accuracy	Has a true representation of the real world been achieved?	Absolute High Plausible (…)	4 3 2
Precision	Is the degree of precision as good as it can be for the phenomenon being measured? Could it be finer? Should it be coarser?	Excellent Good Fair (…)	4 3 2
Peer review			
Extent	How widely reviewed is the process and the outcome?	Wide Moderate Limited (…)	4 3 2
Acceptance	How widely accepted is the result?	Total High Medium (…)	4 3 2
State of the art	What is the degree of peer consensus about the state of the art of the field?	All but cranks All but rebels Competing schools (…)	4 3 2
Validity			
Relevance	How relevant is the result to the problem in hand?	Direct Indirect (…)	4 3
Completeness	How sure are we that the analysis is complete?	Total (…)	4