PROBABILISTIC RISK ASSESSMENT FOR THE SIZEWELL B PWR

T.P. Speed

CSIRO Division of Mathematics and Statistics
Canberra, Australia.

[Editor's Note: This article was sent to the newsletter in a much longer form, and has been abridged. The original article has been passed on to the nuclear group].

1. INTRODUCTION

Several years ago I wrote a short critique of the 1975 Reactor Safety Study (RSS), US Government (1975), also known as WASH-1400 or the Rasmussen Report. As a mathematical statistician, my interest in the RSS focussed on the nature and quality of the data used, the probabilistic and statistical assumptions made, and the methods of analysis adopted in the Study. My intention was to form an opinion on the validity of the probability figures which were the principal conclusions of the Study, and I concluded that they were totally valueless. Not long afterwards a review of the RSS, US Government (1978), concluded, amongst other things (see p.viii) that

> "We are unable to determine whether the absolute probabilities of accident sequences in WASH-1400 are high or low, but we believe that the error bounds on those estimates are, in general, greatly understated".

Now, nearly a decade later, it seems that few people believe the probability figures calculated in WASH-1400, although most would accept that the process of carrying out a probabilistic risk assessment has considerable value. The proposal by the United Kingdom Central Electricity Generating Board (CEGB) to build a pressurized water reactor (PWR) at Sizewell in Suffolk following the Standardized Nuclear Unit Power Plant System designed by Bechtel for Westinghouse in the U.S., and the subsequent production by Westinghouse of the Sizewell B Probabilistic Safety Study (PSS) for the proposed reactor, offer us an opportunity to see the extent to which the problems identified in the RSS have been overcome. Like the RSS, the Sizewell B PSS aims for, and obtains, figures purporting to quantify the probabilities (per year) of core melt and release from the containment, although there are no error bounds of any kind attached to these figures.

The main aim of this paper is to examine the Sizewell B PSS in much the same way as I did the RSS. I sought evidence of careful data analysis, an appreciation of the difficulties and uncertainties inherent in the task, perhaps consideration of robustness issues and other modern methods but at least a critical use of traditional techniques; in short, evidence that the probabilistic and statistical analyses of the Study were carried out by competent professionals in the area.

2. THE SIZEWELL B PROBABILISTIC SAFETY STUDY

(i) Background and structure of the Study.

In June 1981 a Task Force was set up to develop firm design proposals for the pressurized water reactor which the United Kingdom Central Electricity Generating Board (CEGB) wished to build at Sizewell in Suffolk, see Marshall (1983). In that article Marshall gives an outline of the design process followed in the U.K. leading up to the Statement of Case, Reference Design and Pre-Construction Safety Report published by the CEGB in May 1982. As noted earlier in this paper he states that ".... The target aimed for is a risk factor of less than once in a million years." At around the same time the Sizewell B Probabilistic Safety Study, Westinghouse (1982) appeared, probably in anticipation of the public enquiry into the building of the reactor, showing that the target had been achieved.

The overall conclusions of this Study are

> "that the frequency of core melt at the Sizewell B plant from internal initiators is conservatively estimated to be 1.16×10^{-6} per year and the release frequency from the containment is 2.8×10^{-8} per year. These releases are largely dominated by either basement failures which are not as serious as above ground failures or delayed overpressure failures which would not occur for at least 8 hours after accident initiation."

How did the Sizewell B PSS reach these conslusions? Although structured differently, the basic approach adopted in the Study was the same as that of WASH-1400, namely (p.1.1-1)

> "Component failure distributions are determined. Equipment reliability is calculated. Physical phenomena are evaluated. Finally, the probability of any accident with any number of safety system failures can be calculated along with the associated release of radioactive material."

This then is the structure of the probability calculations undertaken. A question which is not addressed anywhere in the Study is: What evidence is there that probabilities calculated using this model bear any relation to the behaviour of real-world PWRs and reactor accidents?

(ii) Types of probability evaluations used.

As with WASH-1400, the Sizewell B PSS contains an undifferentiated mixture of probability evaluations based on hard data, apparently quite subjective considerations, models of physical phenomena, and other methods. The conslusions of the Study are that "the frequency of core melt .... is conservatively estimated to be ...." and it seems worthwhile to note some of these evaluations in order that the reader can focus better on the question just mentioned (which is not addressed in the Study).

Apart from the probabilities concerning the failure or unavailability of equipment, the main data-based probabilities in the Study concern the initiating events and I comment on these in (iii) below. In Appendix 3.8.1 entitled Hydrogen Mixing Probability Assignment we find that the probability of a detonation event in the context of hydrogen mixing is assigned epsilon, with epsilon stated in a footnote to equal 10^{-4}. Further down a value of 10 epsilon is used to define a related probability "in order to show that it is greater than" the former, and finally, because two "separate independent constraints must be met" the "probability of a detonation is somewhere between 10 epsilon squared, and epsilon squared, ... In other words, ... small enough that the possibility of a detonation need not be considered in the hydrogen burn question."

This sort of argument is reminiscent of WASH-1400 at its worst, and suggests that not all the critisisms made of that report have been understood by the writers of the PSS. The lapse is not an isolated one, for the following section headed "Probability assignment" apparently quite arbitrarily assigns probabilities of 1-epsilon, 0.95, 0.99, 0.6, 0.7, 0.95, 0.99, 0.6, 0.4 and 0.6 to a string of events associated with hydrogen mixing. When one consideres the anxiety and uncertainty associated with the hydrogen bubble at Three-Mile Island, U.S. Government (1979), the discussion just referred to seems far from adequate.

(iii) Data and data analyses performed.

The data used in the Sizewell B PSS concerns initiating events, hardware (components) and human factors. I will comment briefly on each of the last two below, my main remards being directed at the Study's use of the data embodied in Tables 1.2-13 and 1.2-14. It is on the basis of data in these tables that the Study team evaluated the probabilities ϕ_i of initiating events and their methods of doing so seem to be at least as arbitrary as some of those used in WASH-1400.

On what data are the PSS probabilities based? This is not entirely clear, for although they reproduce a 30x14 table of population event data (not shown), 30 PWRs with a combined operating experience of 131 years and 14 types of events, their later calculations use figures described in Table 1.2-15 as Sizewell B Type Plant Specific.

We quote from p.1.2-5

> "In recognition that the Sizewell B plant has no operating history to extract plant specific data a "like" plant was substituted. The selection of a "like" plant was necessary, since generic plant data would cover too broad a spectrum and vintage of plants (i.e., two, three, and four loop plants, etc.). The lack of a sufficient four loop plant specific data base prevented a generic four loop approach. The Zion plants of the Commonwealth Edison Company were chosen as the like plant from which to extract plant specific data."

It would seem, then, that probability evaluations relating to the Sizewell B PWR should be based upon the experience of the Zion plants.

What can we make of the population data? Denoting by λ_{ir} and n_{ir} the rate and observed number of initiating events of type i and by t_r the number of operational years associated with PWR r, i=1,...,14, r=1,...,30, we can consider as plausible the model which has the n_{ir} independently distributed as Poisson variates with expectation $\mathbb{E}(n_{ir})=\lambda_{ir}t_r$. One reasonable interpretation of the term population would be that the λ_{ir} do not depend on r, i.e. are homogeneous:

$$\mathbb{E}(n_{ir}) = \lambda_i t_r \qquad (H)$$

and one can readily fit and test this model. Indeed one can also fit and test the multiplicative model

$$\mathbb{E}(n_{ir}) = \alpha_i \beta_r t_r \qquad (M)$$

with some convention identifying the parameters α_i and β_r. These models are readily fitted (omitting the four events with no occurrences to date), and the likelihood-ratio (approximate) chi-square test statistics for them are 527 on 261 d.f. for (M) and 996 on 290 d.f. for (H). Clearly neither is a plausible model for the data. How then should we summarise it? One could, preferably with further information, group the reactors into more homogeneous subsets, and pool across these, perhaps assigning the probabilities of one of these groups to the proposed Sizewell B reactor. Alternatively one could simply leave the population as it is, and bear in mind that any pooled estimates of rates (probabilities) would have greater than the usual (Poisson) uncertainty associated with them.

The Study does neither of these things. It presents Table 1.2-15 (not shown), having given in the text the following almost incomprehensible explanation (pp.1.2-6, 1.2-7):

"The plant-specific and the population data is summarized in Table 1.2-13 by plant and by initiating event. Table 1.2-14 shows the number of operation years in the data base by plant.

It should be noted that these frequencies are derived using Bayesian techniques (1-3) and represent a conservative approach to initiating event qualification. This is particularly true for rare events. Rare events are defined as initiators which have not occurred within the ∿200 years of PWR operating experience.

Table 1.2-15 shows the probability of occurrence of each initiating event expressed in terms of various lognormal parameters for the plant-specific and PWR population generic data. A review of the Sizewell B design against typical PWR design and review of consequential failures was performed to identify inconsistencies within the data base. This review indicated that the small LOCA frequency as an initiator was very conservative, most small LOCAs have occurred as consequential failures (pump seal due to loss of cooling or PORVs sticking when challenged). Since these consequential failures are explicitely treated via looping functions as described in Section 1.6.2, the small LOCA frequency as an initiator, alone, is set consistent with the medium and large LOCAs."

Let us look closely at what these (presumably) Bayesian techniques, the use of the log-normal distribution and review of the Sizewell B design have done with the data. We will compare them with results from a classical statistical analysis, assuming a Poisson process for the events and homogeneity for the population data; see Table 1 for selected events. Incorporating the evident inhomogeneity would have little effect on our estimated rates but would increase their standard errors.

There are several points to note in connection with our Table 1. Firstly, the obvious classical estimates for the rate of small loss of coolant accident (LOCA) are either 1 in 10 (specific) or 1 in 50 (population); the PSS uses 1 in 1000, and asserts this with a degree of confidence whose upper 97½% point is still 1 in 200! The same figures are used for Large LOCAs, whereas the upper limit of a classical 95% confidence interval for this rate, based on no events observed, is either 1 in 3 (specific) or 1 in 40 (population). Secondly, we observe that there is some consistency between the two approaches for Turbine Trips (specific), although a big difference for the population rates. And here the numbers are quite large. It is apparent that some very unusual statistical methods are being used. Thirdly, we note that in 10 out of the 13 events whose probabilities are evaluated in the PSS Table 1.2-15, the precision claimed for the plant specific estimate is greater than that claimed for the population generic, a quite impossible conclusion to draw from data alone. Clearly the review of the Sizewell B design has produced these figures, not any statistical analysis, Bayesian or otherwise.

Estimation of probabilities of certain initiating events

				Classical Statistics		*Sizewell B PSS*	
E	*S*	*n*	*T*	*MLE*	*95% Conf. Int.*	*Mean*	*95% Interval*
(2)	P	3	131	$22{\cdot}9\times10^{-3}$	$[4{\cdot}6,67{\cdot}2]\times10^{-3}$	1×10^{-3}	$[{\cdot}02,5{\cdot}9]\times10^{-3}$
	S	1	11	$90{\cdot}9\times10^{-3}$	$[9{\cdot}1,509]\times10^{-3}$	1×10^{-3}	$[{\cdot}02,5{\cdot}5]\times10^{-3}$
(1)	P	0	131	0·000	$[0{\cdot}0,28{\cdot}2]\times10^{-3}$	1×10^{-3}	$[{\cdot}02,5{\cdot}9]\times10^{-3}$
	S	0	11	0·000	$[0{\cdot}0,336{\cdot}4]\times10^{-3}$	1×10^{-3}	$[{\cdot}02,5{\cdot}5]\times10^{-3}$
(11)	P	371	131	2·83	[2·54,3·12]	4·00	[0·66,13·43]
	S	41	11	3·73	[2·67,5·05]	3·69	[2·70,4·92]

TABLE 1

Key:

E: Events (2): Small LOCA; (1): Large LOCA; (11): Turbine Trip

S: Source of data P: Population generic; S: Sizewell specific = Zion 1 + Zion 2.

n: number of events.

T: number of operating years.

MLE: Maximum Likelihood Estimate

95% Confidence interval : see below.

From Johnson and Leone (1977,p.525) the 95% confidence interval for a Poisson rate with n events observed in T units of time is:

n	*Lower*	*Upper*	*n*	*Lower*	*Upper*
0	0	3·7/T	3	0·6/T	8·8/T
1	0·1/T	5·6/T	41	29·4/T	55·6/T

When n>50 a normal approximation was used.

For lognormal distribution with mean $\alpha = \exp(\mu+\frac{1}{2}\sigma^2)$ and variance $\beta = \alpha^2[\exp(\sigma^2)-1]$, the median is $\exp(\mu)$ and scale factors giving the 2½% and 97½% probability points are

$$\exp \mp [1{\cdot}96\{\log\{\frac{\beta}{\alpha^2} + 1)\}^{\frac{1}{2}}] .$$

The reliability data on components is stated as being from Masarik (1981), which I have been unable to consult, other (unspecified) sources, or engineering judgement. It is still quite disconcerting for a statistician to come across the following: in Table 1.3-1 headed Component Failure Data (p.1.3-6)

Manual valve, normally closed : ϵ

Elsewhere (p.1.5-8) we find the same, namely:

> Assumption 3: All manual valves in an in-service train will have negligible (ϵ) failure probability.

The value of ϵ is not specified in this context but elsewhere, as we have already noted, ϵ is taken as 10^{-4}. However, there are other probabilities of this magnitude or smaller in Table 1.3-1.

(iv) Multiplication of probabilities, including treatment of common-mode failures.

The general form of the multiplication rule for probabilities is

$$pr(A\&B\&C\&...|H) = pr(A|H)pr(B|A\&H)pr(C|A\&B\&H)... \qquad (7)$$

and in the present context, a large number of such probabilities of conjunctions of many terms are being calculated. To what extent is (7) fully appreciated and used in the Sizewell B PSS? In my opinion it is probably not properly understood. There are so many tacit probabilistic independence assumptions and inconsistencies involving (7) that an explicit concern for common-mode failures (see below) seems quite beside the point.

Consider the basic structure of the calculation of the probability $pr(R_k)$ of release category k. It expresses the answer as a sum of terms each should be of the form $pr(E_i)pr(D_j|E_i)pr(R_k|D_j\&E_i)$, where E_i is initiating event i and D_j is plant damage state j. i.e. it assumes that for all i,j and k we have

$$pr(R_k|D_j\&E_i) = pr(R_k|D_j). \qquad (8)$$

Expressed another way, (8) states that release events are probabilistically independent of initiating events given any of the plant damage states for which (8) is non-zero. This might sound plausible when said in isolation but it seems to contradict quite explicitly the discussion in 2.2 Selection of Key Accident Sequences. There we find the initiating events playing a role (see esp. p.2.2-4 and Table 2.2-3) in the evaluation of probabilities associated with the various resease categories, although these have not yet been defined at this point in the Study. Thus the basic structure of the Study calculation appears to be flawed because of a lack of appreciation of (7).

How much credibility can we place in probabilities calculated as products of 8, 10, 12 or more terms, assuming independence all along the path? In the PSS figures of the order 10^{-15}, 10^{-19} are obtained in this way. I would suggest that there is little or no body of reliability experience with such complex systems - I know of none - in which this sort of calculation has led to results which related reasonably closely to subsequent experience. Is it not reasonable to expect the PSS to convince us that these methods will be effective in their case? Or should we be content with the knowledge that they tried as hard as they could to bridge the gaps in their knowledge?

The issue of common-mode failure probabilities - how do we calculate pr(A&B|H) when we only know pr(A|H) and pr(B|H)? - is raised in this Study as it was in WASH-1400. Of course all that can be said without extra information is

$$\max(1-pr(A|H)-pr(B|H),0) \leq pr(A\&B|H) \leq \min(pr(A|H),pr(B|H)).$$

The Sizewell B rejects the thoroughly discredited WASH-1400 "solution" to the "problem" of common-mode failures and embraces a different false technique: the additive cut-off approach. From p.1.3-3 we quote:

> "The additive cut-off approach limits the system unavailability to a pre-determined minimum (i.e., "cut-off") value. However, since there is no statistical basis for the choice of the cut-off value, it must be selected entirely on engineering judgment.
>
> The choice of the cut-off value is no easier than the choice of the approach. The CEGB Design Safety Guidelines require that the CMF probability shall not be assumed to be less than 1.(-5) and the system unreliability shall not exceed 1.(-3) overall.
>
> Thus a value of 1.(-4) per demand was chosen for the CMF cut-off contribution for all systems, with the exceptions noted below."
>
> The arbitrariness of the procedure is quite evident.

Finally, we observe that no such "additive cut-off" was used in the probability multiplications carried out on the various event trees. Surely the case for the use of such a rule there would be stronger; 10^{-4} instead of 10^{-19} in such examples would certainly be more worthy of the description conservative. It is ironic that in such calculations a probability of zero is distinguished from one of 10^{-17}, 10^{-18} or 10^{-19}.

(v) Completeness of the Study

It is possible that the members of the Study group identified all accident sequences which could contribute significantly to the risk? The discussion in the RARG Section III is relevant here, although essentially leaving the question unanswered, but in a footnote (p.15) we find the opinion:

> "One of us (F. v. H.) questions whether, for a system as complex as a nuclear power plant, the methodology can be implemented to give such a high level of confidence that the summed probability of many known and unknown accident sequences leading to an end point such as a core melt is well below the lomit set by experience."

Experience with actual reactor accidents (Browns Ferry, Three-Mile Island) would seem to support this view and it appears to be shared by Critchley (1976,p.18):

> "No high-risk, major-hazard, safety-assured plant like a nuclear reactor should be built unless it is so well designed, constructed and operated that disastrous failure cannot be foreseen in the anticipated circumstances of its existence; that is, such an event must be 'incredible'. Thus, the permitted net chance of occurrence of a catastrophic radiation accident arising from any envisaged cause must tend to be vanishingly small. A risk so forecast cannot be true. The true hazard is given by the summation of the occurrence probabilities of all accident-producing causes which includes an almost infinite spectrum of unexpected, unusual or highly improbably though possible happenings or coincidences. At the present time, at least, the task of catching such a large number of rare, random and diverse things is Sisyphean. There is, thus, a severe limitation on the input data which vitiates any quantitative predictions, and such serious accidents as might occur will be most likely to be 'rogue' events which would not be identified in the quantifier's philosophy."

(vi) Sensitivity analyses carried out.

It is common in the discussion of a complex mathematical model for which any check on its realism and predictive power seems out of the question to vary some of its parameters or assumptions and note the effect on certain overall features of the model. Exactly what we learn about the validity or usefulness of a model by doing this is far from clear, but such a practice can certainly help to isolate features of the model which might otherwise escape notice.

A very perfunctory sensitivity analysis of the kind just described was carried out in the Sizewell B PSS. For example, the assumption of quarterly testing of sump recirculation valves is changed to annual testing and in this case the base figure 1.16×10^{-6} changes to 2.7×10^{-6}, at the same time, apparently, decreasing the frequency of two types of plant damage states. Just how this theoretically impossible conclusion resulted is difficult to discover.

3. CONCLUSIONS

Having read a good deal of the Sizewell B PSS fairly carefully, and having found clear evidence that the Study group failed to appreciate a number of important points of probability and statistics, that they failed to use clear and accepted methods of statistical analysis with their data, having seen no evidence of worthwhile advances on problems in the area well known to be difficult, if not insurmountable, I find myself concluding that there is no reason at all for me to accept their final probability figures as having any value. Many of the reasons which have led me to this conclusion have been detailed above; many others have been omitted.

Do I think that with careful, competent statistical analysis this approach could yield probability figures which I might believe? The answer is angain no, I do not, for much the same reasons as those put forward by O.H. Critchley in the extract quoted above. I think that at many important points in the analysis of accidents, probability methods are quite irrelevant and their use might give a dangerous sense of security to analysts. The WASH-1400 discussion of the Browns Ferry fire illustrates this point very well.

The fact that the groups who write reports such as the Sizewell B PSS or WASH-1400 invariably show such a lack of appreciation of the subtleties of probability and statistics I see as initimately connected with their obvious belief and hope that the approach will yield worthwhile probability figures; if they understood the subject far better, they would expect far less from it. If actual figures are really required, workers in the field of reactor safety should turn to the calculation of well-based and realistic risk assessments which, although they might not be as low, would be believable and should also be closer to the truth. There are many statisticians who would be willing and able to collaborate in such an exercise.

REFERENCES

CRITCHLEY, O.H. (1976) "Risk prediction, safety analysis and quantitative probability methods - a caveat". J. Br. Nucl. Energy Soc., 15, 18-20.

MARSHALL, SIR WALTER (1983) "Design and safety of the Sizewell pressurized water reactor". Proc. Roy. Soc. Lond. A, 385, 241-251.

MASARIK, R.J. (1981) "Selected failure-rate data for Westinghouse NSS Components". Westinghouse Electric Corporation. [Cited in WESTINGHOUSE ELECTRIC CORPORATION (1982)]

PSS : See WESTINGHOUSE ELECTRIC CORPORATION (1982).

RARG : See US GOVERNMENT (1978).

RSS : See U.S. GOVERNMENT (1975).

U.S. GOVERNMENT (1975) Reactor Safety Study WASH-1400 (NUREG75/014), Washington: NRC. [Also known as the Rasmussen Report.]

U.S. GOVERNMENT (1978) Risk Assessment Review Group Report to the U.S. Nuclear Regulatory Commission. NUREG/CR-0400. Washington : NRC. [Also known as the Lewis Report].

U.S. GOVERNMENT (1979) The Need for Change: The Legacy of TMI. Report on The President's Commission On The Accident at Three-Mile Island. New York : Pergamon Press.

WESTINGHOUSE ELECTRIC CORPORATION (1982) Sizewell B Probabilistic Safety Study. Westinghouse Electric Corporation WCAP 9991 Rev. 1. [Referred to in text as Sizewell B PSS or just PSS.]