April 19, 1981

Comments re the Vallecitos Nuclear Reactor / GETR

David R. Brillinger
Department of Statistics
The University of California, Berkeley

Preamble. I was requested by an Aide to Congressman R. Dellums to review certain materials relating to the siting of nuclear reactors at Vallecitos, California.

This material was:

1. Probability Analysis of Surface Rupture Offset Beneath Reactor Building General Electric Test Reactor (12 April 1979) EDAC-117-217.13
2. Additional Probability Analyses of Surface Rupture Offset Beneath Reactor Building General Electric Test Reactor (March 12,1980) JBA-111-013-01
3. Probability Analysis for Combined Surface Rupture Offset and Vibratory Ground Motion General Electric Test Reactor (April 29,1980) JBA-111-014-01
4. Combined Parameter Probability Analysis General Electric Test Reactor (July 18, 1980) 111-017-02
5. Letter from L.C. Rouse, NRC to R.A. Koschner, GEC dated Feb. 26, 1981

Brief Summary of Review. The analysis provided in Reports 1 - 4 is incomplete, simplistic and unaccompanied by proper qualifications. I do not believe that any reasonable statistician would let himself be associated with probabilistic and statistical results of the sort provided. The approach does not address the properly skeptical audience. The NRC has not been well served. The conclusions are extremely debatable.

A desire for simplicity is commendable; however simple models can be very missleading, and this may well have happened here. It appears to me that too many simplifying assumptions have been made. The problem has been condensed too far. It has been treated as one-dimensional, instead of its actual three-dimensional nature. It has been treated as static, instead of its actual dynamic nature. Important variables have been omitted. The end result of the analysis, a single number (probability), constitutes too brutal a summary of the situation. Possible damage to the GETR and the human population depends on much more.

The usual scientific procedure of pointing out the limits of uncertainty in the findings and inference has not been followed. Many assumptions made have not been justified and may at best be described as guesses or convenient approximations. A cursory review of the literature yields conflicting physical values for many of those employed. Other values and physical models fit the data equally well and some undoubtedly lead to quite different end numbers.

The letter 5. strikes me as premature. It refers to a deterministic approach and a probabilistic approach. In fact, deterministic approaches are probabilistic. The authors neglect phenomena that have small probability of occurring and employ parameter values subject to measurement error. The probabilistic approach values recorded in the letter constitute too savage a summary of the situation, are based on data subject to error and are based on debatable assumptions. The implications of deliberately building in bias, (conservatism?), need to be investigated.

A full risk study should be carried out and the approach of such a study assessed in detail. The study should include error analysis, critical examination of assumptions, specific study re all reactors at the site, Monte Carlo / simulation work, be multivatiate rather than scalar, and examine ground-to-structure interaction (feedback) among other things. If a decision to resume activities is made, all assumptions and procedures of the analysis should be set down in as specific and detailed fashion as possible, in order that if an accident or other surprise occurs, specific scientific knowledge will be gained (rather than only the knowledge that someone's engineering judgement was incorrect.)

One final comment. It seems that in a probabilistic approach, ratios of probabilities of alternatives are relevant as well as absolute probabilities. In particular, are not the risks associated with other sites clearly less than those of Vallecitos?

Report 1 - EDAC-117-217.3 . Suppose a coin is flipped n times and a tail comes up each time. Are you then willing to proceed on the basis that the probability of a head the next flip is 1/(n+2) ? This is the essence of what the authors of this report would have you do. (See expressions (5-11), (5-12).) The result is known as Laplace's Rule of Succession and was popular in the nineteenth century. It has since fallen into total discredit, except for hypothetical situations in which the precise mathematical conditions for its validity hold, and which are far from true here. The formula comes from adopting a Bayesian approach to the problem - assuming that constants are in fact random variables and indeed, random variables whose distribution is known.

I make two remarks. First, the Bayesian approach is not generally accepted by the vast majority of statisticians (unless certain very specific conditions obtain). It is exceedingly controvercial. Second, quoting the report page 2-1 quoting USNRC Standard Review Plan, Section 2.2.3: "Acco[illegible]gly, a conservative calculation showing that the probability of occurrence of potential exposures in excess of the 10CFR Part 100 guidelines is approximately 10^{-6} per year is acceptable if,..." . In my opinion the results derived from Bayesian arguments, of the sort presented here, ARE NOT IN FACT PROBABILITIES. Computing them bears no relationship to satisfying the quoted requirement.

Other comments I have include:

1. Exceedingly precise assumptions are set down, yet there is so little data made use of, and the situation is so important.
2. The word "conservative" and even "very conservative" is used throughout. Because of the many unjustified assumptions, to my mind, these words are totally inappropriate.
3. The problem is reduced to a one-dimensional one - of a point (the fault entry) being located in a certain interval (the reactor building) of a line. The problem in undeniably three-dimensional. The building has three dimensions; length, width and depth. The building does not run parallel to the fault. In a one-dimensional approach, its widest prospect should be employed. I would argue, that a relevant probability to seek to evaluate is that of a curve (not straight line) intersecting a box of the length, breadth and depth of the situation. The probability actually evaluated is less than this probability, and hence not

conservative.

4. Page 5-1. A Poisson process is assumed. I have analyzed many seismic event series, and I feel read the vast majority of scientific papers written on such analyses. The Poisson assumption simply does not hold. The times are clustered. (Nor is the Poisson assumption off in a conservative direction. Making it results in seemingly more precise estimates.)

5. I sense a belief on the part of the report's authors, that any new movement is a lot more likely to take place on one of the existing shears. If this is the case, why is there more than one shear at the site? (See eg. p 5-7 .)

6. The report takes no note of the (possibly substantial) measurement error in the data. Nor does it model in variation result' g from an (unknown) varying number of offsets.

7. Expression (5-2) corresponds to a probability of exactly one offset. What is required is a probability of at least one (a larger value.)

The report does also provide a non-Bayesian computation, estimating the yearly rate of shearing between the two given shears by the upper limit of a one-sided 95 per cent confidence interval (the data being no shears in 128,000 or 195,000 years) and then approximating P_1 by this value. This approach depend strongly on the formulation adopted and the specific Poisson assumption (which as indicated above is strongly debatable.) No discussion of the estimate's variability is provided. Provision of variability is standard in the classical approach. (I remark that this report, page 7-1, provides the first time in my fairly long career that I have seen a 95 per cent confidence level described as "very conservative".

Within the NRC scheme of allowing the neglect of events with probability approximately 10^{-6}, this report comes no where near such a demonstration. Quite frankly, it is at the level of a first or secong year undergraduate paper.

Report 2 - JBA-111-013-01. This memorandum adresses the same type of problem as the first report. It is more sophisicated and goes into the issues in greater detail, however many of the above criticisms apply to it equally and further criticisms can be made. Two approaches are provided; one Bayesian, one classical. Once again I question whether in fact "probabilitiies" provided by a Bayesian analysis are probabilities in an appropriate sense for NRC regulations. Once again a one-dimensional model is assumed. The three-dimencional model should be dealt with (and an artificial assumption removed from the problem.) Certain distributional assumptions are made. These may be checked with Sieh's data. Why wasn't this done? Reading through the report one notes a lot of assumptions being made. Yet there is minimal, if any, examination of the reasonableness of the assumptions. In the classical approach provided, estimates of probabilities are derived - yet there is no indication of their sampling error. Once again results, that are far from so, are described as "conservative" and even "very conservative". No account is taken of the measurement error and biases in the data analyzed.

Specific comments include:

1. Page 2-3. The independent binomial trials assumption will not be true. The degree of dependence of the results on the assumptions needs to be checked. Expression (2-2) is simply wrong or at a minimum of no use. The authors want to use it with a large value of t^*; however because they set p up as a conditional probability, and because the conditioning event never happened, the expression is vacuous. (The authors shouldn't have so casually (as opposed to conservatively) set $P_{\overline{ON}}$ to 1. .) The expression (2-3) is available for manipulation at will. By choice of C one can get any value one choses. (The authors consider C = .95, .90, .10, but provide no justification.)

2. Page 2-4. Quite a number of very debatable assumptions are made on this page, with insufficient examination. Eg. specific distributional forms, equally likely movement on shears. The uniform distribution DOES NOT produce "maximum conservatism" as claimed.

3. Page 3-1. $\sigma/\bar{t}$ = .5 is an estimate. What about its sampling fluctuations?

4. Page 3-2. Sampling fluctuations and measurement errors not taken account of in these procedures.

5. Table 3-4. Sieh(1978) J. Geophysical Res. 83, 3907-3939 indicates that the dates are subject to substantial measurement error (indications provided) and that more ruptures may have taken place. Further, the last value is given by him as 545, not the 575 of the Table.

6. Page 4-1. Note all the assumptions made. How much faith can one have in the resulting values ?

7. Page 4-2. Maximum likelihood is used for estimation. This procedure has known difficulties/requirements (eg. large samples for optimality, regularity conditions.) These are not commented on in this very small sample situation.

8. Page 4-3. That the Central Limit Theorem is appropriate for $N > 5$ is very debatable. Simply assuming a distribution for N, as is done here, has to be justified.

9. I regard the numerical values determined in this Approach 1 as totally unjustified and as based on a procedure subject to easy manipulation. The authors are to be commended for seeking out data. However, why should San Andreas results be relevant? The fault is so very different.

10. Approach 2. This is Bayesian. Highly debatable. And in the end essentially comes down to nothing more than Laplace's Rule once again. I do not see how any important decisions could be based on its results - further I do not think its use should be accepted as meeting NRC probabilistic risk requirements.

In summary, the authors of this report made many highly specific assumptions and yet spend little time justifying the assumptions. Many of the assumptions are wrong. Certain of the statistical procedures employed are inappropriate. It is hard to see what meaning, if any, the numerical values produced have. GE's March 12, 1980 letter of transmittal describes the assumptions made as conservative and the estimate provided as "best". This seems hardly to be the case. Evidence that the probability is of order 10^{-6} has not been provided.

<u>Report 3 - JBA-111-014-01</u>. This report continues the one-dimensional treatment of the geometry of the problem, and hence is not conservative. In the derivation of the probabilities, certain implicit assumptions are made re the character of the offset taking place. The analysis provided is static, rather than the actual dynamic. (Apparently the acceleration values will be used later to scale certain spectra; however a single number cannot come close to dealing with the complexity of the problem.) The probability levels that were derived in Reports 1, 2 are used to motivate choices in this report. Tension and friction between the soil and the foundation are not discussed.

The lognormal distribution is used to describe acceleration values. My review of the literature suggests that its use has not been scientifically justified. The effect of the actual distribution being something else has hence to be studied. Further the parameter values of the distribution (.3g, $2^{1/2}$) are simply set down. These choices have to be justified. My review of the literature turned up lots of larger values, and if anything an indication of higher values being recorded and considered as time goes on.

The intention of the study seems to be to provide pairs of values (acceleration, cantilever length) for use in later studies. A product probability criterion is employed. The best criterion to use depends on what is done later. My feeling is that no criterion should be used at this stage. The separate values should be retained and propagated through the later analyses.

The last expression on page 2-3 contains three typos.

Report 111-C17-02. The general comments re the previous report apply here as well. A new pair of values (maximum acceleration, average cantilever length) is introduced for later use. The average value has clear deffects in the description of a random variable. The separate random values should be propagated through later analyses, as recommended above.

Page 1-1 contains a strange piece of logic. It is argued that it is conservative to neglect certain cases (non-cantilever.)

Whether the results obtained in these last two reports are effective or not depends on the specific uses they will be put to.

Letter 5. The letter brings home the fact that the logical bases, on which the decisions are made, are fluid. Sometimes a probabilistic approach is taken, sometimes a so-called deterministic approach. Deterministic approaches are based on decsisions re likelihood, and values with measurement error. These need to be taken note of in some formal fashion. Further consideration of the deterministic approach brings home the fact (also true for the probabilistic approach) that there is no way to prove that all the (hazardous) possibilities have been considered.

The letter mentions two physical parameters necessary for consideration in making these decisions (acceleration and displacement.) Two parameters are far to few to describe these complicated matters. The wish seems to be to avoid a full analysis, rather to carry through a number of separate self-contained analyses. Simplicity of approach is commendable, however these are serious matters. There are feed-backs and interactions between the parts of the problem here.

E X H I B I T B