APPENDIX I. PRESENTATIONS AT THE MEETING

(5) Accounting for Data Uncertainties in Comparing Risks from Energy Systems, Ulrich Hauptmanns, Universität Magdeburg, Germany
Accounting for data uncertainties in comparing risks from energy systems


Abstract

Data and models for risk comparisons are uncertain and this is true all the more the larger the time horizon contemplated. Statistical methods are presented for dealing with data uncertainties thus providing a broader foundation for decisions. Nevertheless, it has to be borne in mind that no method exists to account for the „unforeseeable“ which is always present in decision making with respect to the far future.

1 Introduction

The debate on the risks of the peaceful use of atomic energy has intensified public awareness of impacts on the environment. It is increasingly understood that almost all human activities may adversely affect the environment, at least if they are carried out on a large scale. This is true even for energy conversion options which for a long time have been considered free from risk, e.g. the use of solar or wind energy. Therefore, it is reasonable to determine the impacts resulting from the use of different energy sources. They should be compared with one another and accounted for when making decisions on how to satisfy the energy demand of a country. It must be kept in mind, however, that risk is only one aspect among many others in the assessment of an energy supply system. According to [1] the following criteria should be considered:

- economy,
- international compatibility,
- environmental compatibility,
- societal compatibility.

The interrelations between the different aspects should be accounted for. This is, however, impossible given the present state of knowledge, particularly, if the degree of fulfilment of the individual criteria is to be quantified.
A prerequisite for comparing risks is the use of clearly defined notions [2]. In addition, the choice of the basis of comparison requires special attention. For example, a power station of 1,000 MWe may serve as a reference; the corresponding number of units would have to be taken if the plants were smaller. A different choice may be the total energy generation of a plant during its lifetime. Another could be the services rendered to the end user; still another the total risk of an entire electrical power grid. In this case the different areas of application of the various processes for base, medium, and peak load would be accounted for. It is evident that the choice of the basis of comparison will influence the calculated risk.

If the far future is the concern, systems have to be examined whose technical maturity has not yet been reached, as for example the use of heliostats or of satellites for converting solar energy. Judgments on the risks of long term storage of wastes have to be made. New, still unknown options may arise. Naturally, in these cases less confidence should be placed on the data than in the case of well-established systems like coal-fired or LWR power stations.

In performing risk comparisons for decision making it must be borne in mind that not all of the available systems may be used in every country. For example, solar heat for electricity generation is not economic in the FRG, because of climatic reasons.

The environmental impact frequently used as a yardstick for comparison constitutes a risk to the population not directly concerned with the energy conversion process. In addition, those directly involved in energy conversion are exposed to specific risks. Among them are accidents which may occur during the mining of raw materials, transport, in the power station itself, or during reprocessing and waste disposal. Occupational diseases may result from adverse conditions at the place of work, for instance, silicosis among coal miners. In addition, the construction of plants may cause accidents. In this context it must be borne in mind that in general the volume of construction increases as the energy density of the exploited energy source decreases. For this reason, the construction of the plants together with the required materials represents an important risk contribution for renewable energy sources like solar or wind energy.

2 Calculational procedures for risk comparisons

The total risk of a process for energy production comprises the risk of the conversion process, a possibly required fuel cycle (including long term storage of wastes, if necessary), and the risk derived from the construction of the plants and from the acquisition of the necessary materials. Other aspects, for example, the risk of an insufficient energy supply as a consequence of renouncing the building of power stations, is not normally addressed.

The assessment of any risk of mass phenomena, for example, from occupational or transport accidents, is based on statistical evaluations of records on the past; for rare events, like core melt in a nuclear reactor, analytical methods are used. Damage to health as a consequence of emissions is estimated by epidemiological investigations and extrapolations from higher to lower doses. In calculating the risk contributions from the installation of plants two different methods are used: process analysis and input-output analysis.

In a process analysis the most important materials for building the plant are identified and their necessary quantities are calculated. Drawing upon statistics for the industrial branches
concerned, the risk caused by their production is assessed. This is augmented by the risks from accidents during the transport of the building materials and the construction. The procedure is schematically shown in Fig. 1 (cf. [3]).

In an input-output analysis the interactions of all segments of an economy are represented in the form of a matrix. The lines contain the sales of the respective segments to all others; the columns state the input of goods and services to one segment of the economy from all the other segments. In this way, all goods and services received from any of the segments of the economy are registered in monetary units. Using appropriate factors these may be converted into risk numbers; for example, accidents per unit quantity of product for the different segments of the economy.

The application of the input-output analysis to the assessment of risks is outlined in Fig. 2, which shows the model developed by the Brookhaven National Laboratory (BNL) for the calculation of the environmental impact and health effects of different technologies. In this model the economy is divided into 110 segments.

![Diagram of a process analysis](image)

**Fig. 1. Outline of the calculation of plant construction risks using a process analysis**

The advantage of using the input-output analysis stems from the fact that the contribution of the entire economy to the construction of a plant is accounted for. Therefore there is no need to determine which materials and activities contribute substantially to risk, as is required for process analysis.

Nevertheless, a limitation of both procedures is that all changes in the economy are assumed to be marginal. The activities of building the power station must not interfere with the economy so as to cause significant changes in the production and accident structure in the segments involved. However, this does not apply to major programmes for promoting certain energy systems. In such cases new production lines will have to be built and the figures from the past will no longer be applicable. A detailed analysis of the economy would then be required and the risk data to be used would have to be estimated by experts.
For treating entire energy systems the risk numbers thus obtained have to be introduced into models where the projections of future energy demand would be combined with risk and possibly economic considerations (cf. [5]). The objective function may be cost minimization using risk as a constraint or vice versa. A serious problem then is that quantities not readily comparable as e.g. possible contribution to the greenhouse effect on one hand and the potential impact of a nuclear accident on the other have to be compared. A way out of the dilemma may be the formulation of levels of aspiration, i.e. setting upper limits to emission or converting the quantities into acceptable societal costs.

The discussion may suffice to show that risk comparisons are beset by considerable uncertainties referring to both models and input data. In what follows only the treatment of data uncertainties will be dealt with. This does not imply, however, that modelling uncertainties should not be addressed.

3 Data requirements and uncertainties

Quantification is only possible if input data is available. They concern the areas indicated in Figs. 1 and 2.

It goes without saying that all this data is uncertain and becomes increasingly so the further we look into the future, because extrapolations loose validity, totally new technical options may become available, new insights into impacts may be gained etc.
Among the sources of parameter uncertainties figure the following:

- difficulties in delimitation
- lack of clarity of statistics
- insufficient knowledge of cause-relationship (e.g. impact of toxic substances, greenhouse effect)
- ignorance about the future

Uncertainty manifests itself in the co-existence of several different data for the same parameter, each of which potentially applies to the problem under study, ignorance in the absence of data.

The inclusion of uncertainties is desirable because it means that all pertinent information is taken into account.

In what follows a possible approach to dealing with data uncertainties is presented.

**4 Mathematical treatment of uncertainties**

The parameters input into the models are normally treated as fixed quantities. In order to account for uncertainties a different statistical viewpoint is taken. The quantities are no longer considered to be fixed but their behaviour is assumed to be random due to one or several of the causes of uncertainty mentioned previously. Random variables are treated by statistical distributions. In this context several cases may be distinguished, each requiring a specific procedure.

1. several values are available
2. two values are available
3. one value and a generic statement on uncertainty is available (e.g. uncertainties of the number of occupational deaths is 10%)
4. only one value is available

The simplest situation is that where two or more observations are available (cases 1 and 2). Then maximum-likelihood estimators may be calculated.

In what follows the procedure is shown for the log-normal distribution. The probability density function (pdf) of that distribution is given by
\[ f(x) = \frac{1}{\sqrt{2\pi \sigma x}} \cdot \exp \left( -\frac{(\ln x - \mu)^2}{2\sigma^2} \right) \quad x > 0 \] (1)

where \( x \) is the random variable in question, e.g., the number dead in the population associated with electricity generation from nuclear fission, \( \mu \) the mean value of the logarithms of that variable and \( \sigma \) the corresponding standard deviation. The distribution is characterized by its expected value

\[ E[X] = \exp(\mu + \frac{\sigma^2}{2}) \] (2)

its median or 50th centile (50% of the values lie below and 50% above)

\[ x_{50} = e^\mu \] (3)

and its 5th and 95th centiles denoted by \( x_{05} \), respectively by \( x_{95} \)

\[ x_{05} = \frac{x_{50}}{K_{95}} \quad \text{and} \quad x_{95} = x_{50} \cdot K_{95} \] (4)

\( K_{95} \) is the uncertainty factor defined as follows:

\[ K_{95} = e^{1.6449 \cdot \sigma} \] (5)

where the value of 1.6449 is the argument of the standard normal distribution corresponding to a probability of 0.95. Hence, as is frequently done, the factor is chosen such that 90% of the values of \( x \) lie between \( x_{05} \) and \( x_{95} \).

The distribution parameters are estimated by

\[ \hat{\mu} = \frac{1}{N} \sum_{n=1}^{N} \ln x_n \] (6)

\[ \hat{\sigma}^2 = \frac{1}{N-1} \sum_{n=1}^{N} (\ln x_n - \hat{\mu})^2 \] (7)

where \( N \) is the number of observations available.

The results of eqs. (6) and (7) are used in eq. (1) to estimate the corresponding parameters there. For the values of Table I we have
Table 1: Occupational deaths from the generation of $2.21 \times 10^{16}$ J of electricity

<table>
<thead>
<tr>
<th>Source</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Q5</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>coal</td>
<td>6.1</td>
<td>4.9</td>
<td>1.6</td>
<td>0.2</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>natural gas</td>
<td>0.1</td>
<td>0.1</td>
<td>0.082</td>
<td>.54</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$\mu_1 = 0.4944 \quad s_1^2 = 1.82$ leading to an expected value of 4.1 and $K_{95} = 9.2$ for coal and

$\mu_2 = -1.9306 \quad s_2^2 = 0.78$ leading to an expected value of 0.21 and $K_{95} = 4.3$ for natural gas.

The density function for coal is shown in Fig. 3.

![Normalized density function for the number of occupational deaths on generating 2.2110^16 J of electrical energy from coal](image)

Fig. 3: Normalized density function for the number of occupational deaths on generating $2.21 \times 10^{16}$ J of electrical energy from coal

Appropriate as well in the present context is the rectangular distribution, whose probability density function (pdf) is as follows.
\[
f(x) = \begin{cases} 
\frac{1}{b-a} & \text{if } b \geq x \geq a \\
0 & \text{otherwise}
\end{cases}
\] (8)

In eq.(8) \(b\) and \(a\) represent the upper and the lower observed value, if only two data are available (case 2) or the greatest and the smallest observation in case of more than two data (case 1), the latter implying that the remainder of the data is ignored.

If only one value and an estimate for the error are given (case 3) the pdf of eq.(8) may be used placing the value in the middle of the error range. The corresponding random quantity is then evaluated from eq.(A2).

Should no error range be indicated (case 4), this may be assessed by an expert. The previously outlined procedure may then be applied using the estimated error range. Alternatively, the estimated error may be expressed in terms of the uncertainty factor \(K_{95}\) of eq.(5) using the observed datum as expected value in eq.(2), whence \(\mu\) can be determined. The random quantity is then calculated according to eq.(A1).

This may suffice for the present treatment keeping in mind nevertheless that other distributions may be more appropriate depending on the nature of the data. However, the procedures outlined here apply mutatis mutandis.

In order to evaluate risks according to procedures outlined in Figs. 1 and 2 addition and multiplication operations have to be performed. The data uncertainties are propagated through these operations using the Monte-Carlo method as described in Appendix A. In doing this, random numbers are used according to the distribution chosen, e.g. log-normal, and a predetermined number of trials, \(P\), is performed.

5 Data referring to the future

Special attention has to be given in the present context to data concerning the future. As far as economic quantities are involved they may be discounted to the same moment in time using standard procedures of financial mathematics.

Risks of technical systems to be compared referring to the future are likely to decrease because it may be assumed that there is progress in safety. The assumption of decrease is, however, beset by larger uncertainties than statements concerning the present.

These larger uncertainties may be accounted for by multiplying the distribution with a correcting distribution represented e.g. by a log-normal with an expected value of 1 and a factor of \(K_{95}\) estimated according to the situation, hence „widening“ the original distribution.

For example, if the data for coal of the Table 1 referred to the year 2100, a decrease of risk of 10\% might be assumed. The uncertainty factor of the correcting distribution is fixed on the
basis of judgment, say $K_{95} = 5$ for the present case. This would lead to a new distribution with a mean value of 3.6 and $K_{95} = 15.5$

Different values adapted to differing time horizons considered are naturally possible. Risk estimates for quite remote options would then be assigned large factors of uncertainty, e.g. $K_{95} = 20$ or $30$.

6 Risk comparisons accounting for uncertainties

If uncertainties are accounted for in comparing the risks of several options, distributions instead of single values have to be compared. The procedure is illustrated by comparing the data for coal with those for gas from Table 1.

The evaluation procedure leads to the convolution integral as described in Appendix B. Its evaluation in the present case gives an overlap probability of 0.066. This is so low that, even in view of the uncertainties involved, the data warrant the statement that the risk for gas is lower than that for coal. This would be true as well, if the corresponding mean values were compared. The situation would change, however, if the data for coal and gas were only known with considerably larger uncertainty, say for the sake of example $K_{95} = 30$. In that case the probability of overlap would be 0.2 and the superiority of gas might well be questioned.

7 Conclusions

The inclusion of uncertainties provides a broader and more reliable basis for risk comparisons. It is mandatory since most data involved are uncertain and the necessity to account for future developments enhances this fact.

Nevertheless it has to be borne in mind that the future also involves unforeseeable developments which cannot be accounted for, however good the treatment of uncertainties.

Appendix A: Monte-Carlo Procedure

The mathematical operations for quantifying risks (e.g. additions and multiplications for the procedures of Figs. 1 and 2) are carried out $P$ times. The results converge stochastically towards the real values with increasing $P$. Each evaluation is a trial which gives a different value for the risk, hence producing a distribution for the risk value, which reflects the uncertainty of the input data. This data is represented by distributions and in every trial a realization in conformity with this distribution is generated, as explained below.

If $Z_{p,1}$ and $Z_{p,2}$ are two independent random numbers uniformly distributed in $[0,1]$ we have for

- log-normally distributed quantities

$$X_p = e^{\sqrt{-2 \cdot \ln Z_{p,1}} \cdot \cos(2 \pi \cdot Z_{p,2})} e^{\mu}$$

(A1)
uniformly distributed quantities

\[ X_p = a + (b - a) \cdot Z_{p,1} \]  \hspace{1cm} (A2)

where \( p \) denotes the pth trial.

**Appendix B: Calculation of the overlap probability of two distributions**

The distribution of the difference of the two random variates defined on the positive half-axis is given by

\[ F_Z = P\{z = y - x \leq 0\} = \int_0^\infty g_Y(y) \cdot dy \cdot dx \]  \hspace{1cm} (B1)

where \( f_X(x) \) and \( g_Y(y) \) both are pdf e.g. according to eq.(1). Evaluation of eq.(B1) leads to the convolution integral

\[ F_Z = P\{z = y - x \leq 0\} = \int_0^\infty f_X(x) \cdot G_Y(z + x) \cdot dx \]  \hspace{1cm} (B2)

The probability of overlap is obtained if \( z \) is equal to 0. Introducing \( f_X(x) \) und \( G_Y(z + x) \) according to eq. (1) then gives

\[ F_Z = \frac{1}{\sqrt{2 \cdot \pi \cdot s_1}} \int_0^\infty \exp \left( -\frac{(\ln x - \mu_1)^2}{2 \cdot s_1^2} \right) \phi \left( \frac{\ln x - \mu_2}{s_2} \right) \cdot dx \]  \hspace{1cm} (B3)

where subscripts 1 and 2 refer to the respective distributions (that for fatalities produced by electricity generation from coal, respectively natural gas in the present case) and \( \phi \) is the standard normal distribution. Eq.(B3) must be evaluated numerically.
References

1. Bericht der Enquete-Kommission "Zukünftige Kernenergie-Politik" Über den Stand der Arbeit und die Ergebnisse gemäß Beschlüß des Deutschen Bundestages - Drucksache 8/2628

2. Hauptmanns, U., and W. Werner:
   Engineering Risks - Evaluation and Valuation Berlin 1991

   IAEA-SM-254/03 (1981)

4. Rowe, M.D.: Assessing systemwide occupational health and safety risks of energy technologies
   IAEA-SM-254/64 (1981)

5. Hauptmanns, U.:
   Risk assessment for interconnected electrical power grids